CAEPR Working Paper
#2007-027

Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital

James Murray
Indiana University Bloomington

November 21, 2007

Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital*

James Murray†
Department of Economics
Indiana University

September 15, 2007

Abstract

This paper examines the empirical significance of learning, a type of adaptive, boundedly rational expectations, in the U.S. economy within the framework of the New Keynesian model. Two popular specifications of the model are estimated: the standard three equation model that does not include capital, and an extended model that allows for endogenous capital accumulation. Estimation results for learning models can be sensitive to the choice for the initial conditions for agents expectations, so four different methods for choosing initial conditions are examined, including jointly estimating the initial conditions with the other parameters of the model. Maximum likelihood results show that learning under all methods for initial conditions lead to very similar predictions as rational expectations, and do not significantly improve the fit the model. The evolution of forecast errors show that the learning models do not out perform the rational expectations model during the run-up of inflation in the 1970s and the subsequent decline in the 1980s, a period of U.S. history which others have suggested learning may play a role. Despite the failure of learning models to better explain the data, analysis of the paths of expectations and structural shocks during the sample show that allowing for learning in the models can lead to different explanations for the data.

Keywords: Learning, firm-specific capital, New Keynesian model, maximum likelihood.
JEL classification: C13, E22, E31, E50.

*I am grateful for the advice and guidance of Eric Leeper, Kim Huynh, Brian Peterson, and Todd Walker; for useful conversations with James Bullard, Troy Davig, Kenneth Kasa, Fabio Milani, Michael Plante, and Bruce Preston; and for comments by the participants of the 2007 Federal Reserve Bank of St. Louis Learning Week Conference, the 2007 Missouri Economics Conference and Indiana University economics department seminars. All errors are my own.

†*Mailing address*: 100 S Woodlawn, Bloomington, IN 47405. *E-mail address*: jmmurray@indiana.edu. *Phone number*: (608)738-5408.

1 Introduction

Rational expectations is one of the most common assumptions in dynamic macroeconomic models. While it is usually made for mathematical convenience, the assumption regarding expectations formation can have non-trivial effects on a model's dynamics. In particular, a large amount of literature has addressed the implications of least squares learning for popular dynamic stochastic general equilibrium (DSGE) models. Agents in a DSGE model that learn do not know the parameters of the model, and instead form expectations by collecting past data and compute least squares forecasts. In this paper I investigate the statistical evidence of learning within the framework of a New Keynesian monetary model and examine the implications of incorporating learning on the predictions of the model

Recent papers have found that least squares learning can important effects on output and inflation determination. Orphanides and Williams (2005b) use an estimated two equation monetary model and demonstrate with simulations of impulse response functions that least squares learning can lead to prolonged inflation following an inflation shock. Using the same model, Orphanides and Williams (2005a) find in another paper that learning on the part of monetary policy can possibly explain the period of stagflation during the 1970s. They suggest that the monetary authority was under-estimating the natural rate of unemployment during this time, and was therefore responding too aggressively to unemployment and not enough to inflation. They suggest that had the central bank responded to inflation instead of unemployment, lower inflation and unemployment would have resulted. Primiceri (2006) suggests that learning on the part of the central bank can explain both the run-up of inflation during the 1970s and the subsequent decline during the 1980s. He suggests that the monetary authority was under-estimating both the natural rate of unemployment and the degree of inflation persistence. Like Orphanides and Williams (2005a), he shows the resulting monetary policy leads to an increase in inflation, but as time progresses the central bank's expectations evolve. The central bank's expectations of the natural rate of unemployment and the degree of inflation persistence return their actual values and therefore the policy prescription becomes stabilizing, resulting in the moderation that occurred from the middle 1980s onward.

The results from these papers depend on a calibrated value for the constant learning gain, a parameter that is responsible for the speed in which expectations evolve, and therefore responsible for the impact learning can have on the dynamics of the model. Milani (2005) is the first paper to estimate the learning gain jointly with the parameters of a model. He finds an estimate for the learning gain which is very close to calibrated values that are popular in the literature. He estimates a standard three equation New Keynesian model and finds evidence in U.S. data that learning explains persistence in output and inflation better than habit formation and inflation indexation. Like the papers cited above, Milani makes specific assumptions about the initial conditions of agents expectations. Many of the initial conditions are set equal to pre-sample ordinary least squares estimates. The exceptions are the degree of inflation persistence, which he assumes is equal to zero, and the sensitivity of output to inflation, which he assumes is higher than the pre-sample evidence.

The results of all of these studies depend on the assumptions for the initial conditions for agents and/or central bank's expectations. These initial conditions are sometimes backed by an economic justification or an argument that such a set of initial conditions accounts well for the data. In this paper, instead of suggesting a specific assumption for the initial expectations of agents, I examine a number of alternative methods for forming these initial conditions. These methods include using the rational expectations solution of the model, using estimates from pre-sample data, and estimating all the initial conditions jointly with the other parameters of the model.

To determine if the estimation results are sensitive to the setup of the New Keynesian model I examine two popular specifications. The first specification is the popular three equation model with the output, inflation, and interest rate determined by utility maximization with the possibility of habit formation, profit maximization under Calvo (1983) sticky prices and inflation indexation, and monetary policy following a Taylor (1993) rule. In this specification there is no capital accumulation and output is produced only with labor. The second specification, suggested by Woodford (2005), extends the model to allow for endogenous capital accumulation, where firms make decisions on investment of firm-specific capital.

There are a number of motivations for extending the empirical analysis to the model with capital accumulation. Including capital in the model introduces data on another variable, aggregate investment, to be included in the estimation procedure. Secondly, introducing capital may alter how expectations are formed, since agents may use past data on capital to make there forecasts. Finally, incorporating capital introduces more expectations into the model which may allow learning may play a bigger role.

The maximum likelihood results indicate that incorporating learning into the New Keynesian models provides little to no improvement in the fit to the data. There is some evidence that the fit improves when using learning frameworks that do not endow agents with data on structural shocks. I find that setting the initial conditions of agents' expectations equal to least squares estimates from pre-sample data actually results in a worse performance of the model. When estimating the initial conditions jointly with the parameters of the model, I find some key parameters governing the persistence and volatility of the model cannot be jointly identified with the initial conditions, suggesting that imposing ad-hock initial conditions can influence the point estimates of these parameters. Finally, plots of the forecast errors, the estimated evolution of structural shocks, and the evolution of agents expectations, are examined to determine if any of the learning models can better explain specific periods of U.S. data, such as the run-up of inflation in the 1970s and subsequent decline. The results indicate that the learning models do not perform any better than rational expectations. Despite the failure learning models to better explain the data, differences do arise in some of the parameter estimates and in the predictions of the structural shocks in the model with capital.

The paper is organized as follows. Section 2 describes the details of the New Keynesian model with and without capital. Section 3 describes the learning process and how learning is incorporated into the model. Section 4 describes the maximum likelihood procedure and the four cases for how initial conditions are constructed. Section 5 reports the results, and section 6 concluded.

2 Model

The New Keynesian model has been used extensively in monetary economics for analysis of theoretical and empirical issues and it is a convenient framework to examine the role of learning on output and inflation determination. Woodford (2003) provides a complete exposition of the model's micro-foundations, its many extensions, and implications for monetary policy, and Appendix A of this paper provides the complete details of the derivation of the model. This paper considers the implications of learning under two popular specifications of the model. The first specification is a framework in which the role of capital is ignored, so that output is produced only from labor. This is a popular specification for empirical work because it allows the model to be written solely in terms of the stationary observable variables: the inflation rate, the interest rate, and the output gap. The output gap is defined as the percentage deviation of real GDP from the value that would occur under full employment and completely flexible prices. The Congressional Budget Office (CBO) provides a measure of the output gap for the United States, which is used when estimating this specification of the model.

The second specification explicitly accounts for endogenous capital accumulation. Output is produced under constant returns using both labor and firm-specific capital. Not only is firm-specific capital a more realistic assumption than a perfect rental market for capital, but Woodford (2005) shows that allowing for firm-specific capital alters the coefficient on marginal cost in the Phillips curve in such a way that allows for greater price flexibility to be consistent with very small values of the coefficient, which is often seen in empirical work, including this paper. The drawback of using this specification is that the model cannot be written in terms of the output gap. The model in this case is written in terms of observable, non-stationary variables: real GDP and real gross private domestic investment. In Section 4 I describe how I account for the issue of non-stationarity.

Both specifications have a continuum of consumers types on the unit interval, and a continuum of intermediate goods producers on the unit interval, each producing a unique intermediate good. Each consumer type possesses a specific labor skill that can only be hired by a corresponding intermediate goods producer. It is assumed that there are many consumers in each consumer type so that consumers do not have market power over the wage. Production of intermediate goods may also depend on capital goods which are firm-specific. Since a capital good in firm i cannot be used by another firm j, there is not a perfect capital rental market which would equalize the marginal product of capital across intermediate good firms. Therefore each firm's labor demand and pricing decision will depend on its current capital stock, which in turn depends on the firm's entire past history.

All the intermediate goods are used to produce a single type of final good, but they are imperfect substitutes for each other in production; therefore intermediate goods producing firms are monopolistically competitive. Prices of intermediate goods are imperfectly flexible according to Calvo's (1983) pricing mechanism where a constant fraction of firms is able to re-optimize its price every period, and the firms selected to do so is randomly determined, independently of firms' histories or characteristics. This setup for sticky prices may seem unrealistic, but Roberts (1995) shows in a model without firm-specific capital that quadratic price adjustment cost, an alternative pricing friction suggested by Rotemberg (1982), yields the same solution as Calvo pricing. The same is not true with firm-specific capital. Under Calvo pricing, at any point in time, each firm will have a different pricing history and therefore a different capital stock. Each firm's relative capital stock will in turn affect the pricing decision. Under quadratic price adjustment costs, all firms face the same friction every period, and so all firms price, labor, and investment decisions remain identical throughout time. Therefore, even though Calvo pricing may seem to be an unrealistic setting, it is a convenient framework to analyze the consequences of the realistic assumption of firm heterogeneity.

2.1 Consumers

Each consumer type has a specific labor skill that can only be hired by a specific intermediate goods producing firm. Since each intermediate goods firm has a different labor demand, wage income will be different for each consumer type. Given a perfect asset market, though, consumption will be equal across all consumers. Each consumer type $i \in (0, 1)$ maximizes utility,

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-\sigma} \xi_t \left(c_t - \eta c_{t-1}\right)^{1-\sigma} - \frac{1}{1+\mu} n_t(i)^{1+\mu} \right],$$

subject to the budget constraint,

$$c_t + b_t(i) = \frac{1 + r_{t-1}}{1 + \pi_t} b_{t-1}(i) + \frac{w_t(i)}{p_t} n_t(i) + \Pi_t - \tau_t$$

where c_t, consumption at time t, is not indexed by individual type i since it is equal across all agents, ξ_t is an aggregate preference shock, $n_t(i)$ and $w_t(i)/p_t$ are the labor supply and real wage of individual i at time t, respectively, $b_t(i)$ is individual i's purchase of real government bonds at time t, r_t is the nominal interest rate paid on government bonds, π_t is the inflation rate, Π_t is the value of profits earned by owning stock in firms, and τ_t is the value of real lump sum taxes. The preference parameters are $\sigma \in (0, \infty)$, which is the inverse of a pseudo intertemporal elasticity of substitution,[1] $\eta \in [0, 1)$, which is the degree of habit formation, and $\mu \in (0, \infty)$ which is the inverse of the elasticity of labor supply. Appendix A shows that

[1] When there is no habit formation, σ is exactly equal to the inverse of the intertemporal elasticity of substitution.

the first order conditions for the consumer lead to the log-linear Euler equation,

$$\hat{\lambda}_t = E_t\hat{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1}, \tag{1}$$

where a hat indicates the percentage deviation of the variable from its steady state.[2] Here, $\hat{\lambda}_t$ is the marginal utility of real income, given by,

$$\hat{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\hat{c}_{t+1} - \sigma(1+\beta\eta^2)\hat{c}_t + \sigma\eta\hat{c}_{t-1}\right] + \left(\hat{\xi}_t - \beta\eta E_t\hat{\xi}_{t+1}\right). \tag{2}$$

I assume that the preference shock, $\hat{\xi}_t$, follows the exogenous autoregressive process,

$$\hat{\xi}_t = \rho_\xi\hat{\xi}_{t-1} + \epsilon_{\xi,t}, \tag{3}$$

where $\epsilon_{\xi,t}$ is independently and identically with mean zero and variance given by σ_ξ^2.

When there is no habit formation, equations (1) and (2) lead to the standard IS equation,

$$\hat{c}_t = E_t\hat{c}_{t+1} - \frac{1}{\sigma}\left(\hat{r}_t - E_t\pi_{t+1}\right) + \hat{\xi}_t.$$

Habit formation is added to the model, because as equation (2) demonstrates, habit formation introduces a source of persistence that does not depend on learning. The larger is the degree of habit formation, the more current period marginal utility depends on past consumption. Since consumption is related to output in the market clearing condition, habit formation creates output persistence. Moreover, Fuhrer (2000) finds that habit formation leads to "hump shaped" impulse response functions, a phenomenon evident in the data.

Appendix A shows that when there is no investment in the model, the demand side of the model can be rewritten in terms of the output gap as follows:

$$\tilde{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\tilde{y}_{t+1} - \sigma(1+\beta\eta^2)\tilde{y}_t + \sigma\eta\tilde{y}_{t-1}\right] \tag{4}$$

$$\tilde{\lambda}_t = E_t\tilde{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1} - r_t^n \tag{5}$$

where $\tilde{y}_t$ denotes the output gap, which is percentage deviation of output from the outcome that would occur under fully flexible prices, and r_t^n is the natural interest rate, the real interest that would occur under fully flexible prices. I suppose the natural interest rate follows the exogenous process,

$$r_t^n = (1-\rho)\,r_n + \rho r_{t-1}^n + \epsilon_{n,t}, \tag{6}$$

where r_n is the steady state real interest rate, and $\epsilon_{n,t}$ is independently and identically

[2] A hat is omitted from inflation because, as Appendix A demonstrates, in order to derive the Phillips curve it is necessary to assume the steady state inflation rate is equal to zero.

distributed with mean zero and variance given by σ_n^2.

2.2 Producers

There is one final good used for consumption and investment which is sold in a perfectly competitive market and produced with a continuum of intermediate goods. The production function is given by,

$$y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di\right]^{\frac{\theta}{\theta-1}} \tag{7}$$

where y_t is the output of the final good, $y_t(i)$ is intermediate good i, and $\theta \in (1, \infty)$ is the elasticity of substitution in production. Profit maximization leads to the demand for each intermediate good,

$$y_t(i) = \left[\frac{p_t(i)}{p_t}\right]^{-\theta} y_t, \tag{8}$$

where $p_t(i)$ is the price of intermediate good i and p_t is the price of the final good. Substituting equation (8) into (7) leads to a consumption price index that holds in equilibrium,

$$p_t = \left[\int_0^1 p_t(i)^{1-\theta} di\right]^{\frac{1}{1-\theta}}. \tag{9}$$

2.2.1 Intermediate goods

The framework for intermediate goods depends on whether the model specification includes capital. In the simple specification that does not include capital, each intermediate good is produced with a unique type of labor according to the linear production function,

$$y_t(i) = z_t n_t(i), \tag{10}$$

where z_t is a stochastic technology shock common to all intermediate goods firms. Output for intermediate good i is demand determined according to the firm's pricing decisions and the demand equation given in (8). For this given level of output, intermediate goods firms choose labor demand to minimize real total cost,

$$C_t(i) = \frac{w_t(i)}{p_t} n_t(i). \tag{11}$$

Appendix A shows when firms hire optimal amounts of labor, the average marginal cost among all the intermediate goods firms (in terms of the percentage deviation from the steady state) is given by,

$$\hat{s}_t = \mu \hat{y}_t - \hat{\lambda}_t - (\mu + 1)\hat{z}_t \tag{12}$$

In the specification with firm-specific capital, the intermediate good is produced with the same labor input and a unique type of capital good according to the constant returns to scale production function,

$$y_t(i) = z_t k_t(i)^\alpha n_t(i)^{1-\alpha} \tag{13}$$

where $k_t(i)$ is capital hired by firm i. For a given level of output, intermediate goods firms choose labor demand and rent capital to minimize real total cost,

$$C_t = \frac{w_t(i)}{p_t} n_t(i) + \rho_t(i) k_t(i), \tag{14}$$

where $\rho_t(i)$ is the rental price of capital good i. Appendix A shows when firms hire optimal amounts of labor and capital, the average marginal cost among all the intermediate goods firms (in terms of the percentage deviation from the steady state) is given by,

$$\hat{s}_t = \frac{\mu+\alpha}{1-\alpha}\hat{y}_t - \frac{\alpha(\mu+1)}{1-\alpha}\hat{k}_t - \hat{\lambda}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t, \tag{15}$$

where $\hat{k}_t$ is the percentage deviation of the aggregate capital stock from its steady state. The technology shock is assumed to follow the exogenous stochastic process,

$$\hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t}, \tag{16}$$

where $\epsilon_{z,t}$ is independently and identically distributed with mean zero and variance given by σ_z^2.

2.2.2 Firm-specific capital goods

Capital goods firms maintain firm-specific capital stocks and rent the capital to the corresponding intermediate goods firm at a real price of $\rho_t(i)$ per unit of capital. This assumption in not essential and is purely used for notational convenience. This model supposes that the market for firm-specific capital is purely competitive, even though firm-specific capital cannot be sold to other firms. This assumption assures an optimal amount of investment in each firm-specific capital good which would be the same outcome if the intermediate goods firms were to invest and own the capital themselves instead of renting it.

Capital goods firms purchase the final good and convert it to a firm-specific capital good. The conversion from a final good to a firm-specific capital good is irreversible and is subject to a stochastic shock, μ_t, that is common to all capital goods. Let $I_t(i)$ denote the purchase of the final good for investment for capital good i, so that $\mu_t I_t(i)$ be the amount a purchase of $I_t(i)$ adds to the capital stock. The evolution of firm-specific capital i is given as,

$$k_{t+1}(i) = (1-\delta)k_t(i) + \mu_t I_t(i) - \frac{\phi}{2}\left[\frac{k_{t+1}(i)}{k_t(i)} - 1\right]^2 k_t(i) \tag{17}$$

where $\delta \in (0,1)$ is the capital depreciation rate and $\phi \in (0,\infty)$ is a capital adjustment cost parameter. When $\phi = 0$, there is no adjustment cost and capital net of depreciation increases by $\mu_t I_t(i)$. Log-linearizing equation (17) then integrating across all the firms leads to the following relationship between capital and investment:

$$\hat{k}_{t+1} = (1-\delta)\hat{k}_t + \delta \hat{I}_t + \delta \hat{\mu}_t \tag{18}$$

Capital goods firms choose investment to maximize the expected utility value of profits,

$$E_0 \sum_{t=0}^{\infty} \beta^t \lambda_t \left[\rho_t(i) k_t(i) - I_t(i)\right], \tag{19}$$

subject to equation (17). The appendix shows that profit maximization leads to the following evolution of the aggregate capital stock:

$$\begin{aligned} &\hat{\lambda}_t + \phi\left(\hat{k}_{t+1} - \hat{k}_t\right) = \beta(1-\delta)E_t\hat{\lambda}_{t+1} \\ &+ \left(\frac{1-\beta(1-\delta)}{1-\alpha}\right)\left[(\mu+1)E_t\hat{y}_{t+1} - (1+\alpha\mu)\hat{k}_{t+1}\right] \\ &+\beta\phi\left(E_t\hat{k}_{t+2} - \hat{k}_{t+1}\right) - \frac{(\mu+1)\left[1-\beta(1-\delta)\right]}{1-\alpha}E_t\hat{z}_{t+1} + \hat{\mu}_t - \beta(1-\delta)E_t\hat{\mu}_{t+1}. \end{aligned} \tag{20}$$

The investment shock is assumed to follow the stochastic process,

$$\hat{\mu}_t = \rho_\mu \hat{\mu}_{t-1} + \epsilon_{\mu,t} \tag{21}$$

where $\epsilon_{\mu,t}$ is independently and identically distributed with mean zero and variance given by σ_μ^2.

2.2.3 Phillips Curve

The Phillips curve is a single equation that describes the relationship between inflation and output, as determined by the supply side of the economy when prices are sticky. The specific price friction employed in this paper is Calvo (1983) pricing. According to this method, only a random subset of intermediate goods firms are able to re-optimize their price in a given period. Allowing for inflation indexation, those firms who are not able to re-optimize their price may adjust their price by a fraction, γ, of the previous period's inflation rate. Let $\omega \in (0,1)$ denote the fraction of firms who are not able to change their prices each period. Since the specific firms able to change their prices each period is randomly determined, ω^T is the probability a firm will not be able to change its price for T consecutive periods. A firm who is able to change its price maximizes the following present discounted utility value of profits earned while the firm is unable to change its price again:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_{t+T}(i)}{p_{t+T}}\right) y_{t+T}(i) - S\left[y_{t+T}(i)\right] \right\}, \tag{22}$$

where $S\left[y_{t+T}(i)\right]$ is the real total cost function of producing $y_{t+T}(i)$ units, given optimal decisions for labor and capital, and $p_{t+T}(i)$ is the firm's price in period $t+T$, given the firm has not yet been able to re-optimize its price. When there is a positive degree of inflation indexation, this price is determined by,

$$\log p_{t+T}(i) = \log p_{t+T-1}(i) + \gamma \pi_{t+T-1} \tag{23}$$

Appendix A shows that the firms' optimal choices for prices in combination with equilibrium in the firm-specific capital goods market leads to the following Phillips curve,

$$\pi_t = \left(\frac{1}{1+\beta\gamma}\right)\left(\gamma\pi_{t-1} + \beta E_t \pi_{t+1} + \kappa \hat{s}_t\right) \tag{24}$$

where κ decreases as ω, the degree of price stickiness, increases. The parameter κ is also a function of other parameters of the model, but in the specification with firm-specific capital, there is not a closed form expression for κ. Appendix A describes the full details of the derivation of the Phillips curve.

When there is no capital accumulation in the model, Appendix A shows the Phillips curve can be rewritten in terms of the output gap as,

$$\pi_t = \left(\frac{1}{1-\beta\gamma}\right)\left[\gamma\pi_{t-1} + \beta E_t \pi_{t+1} + \kappa\left(\mu \tilde{y}_t - \tilde{\lambda}_t\right)\right]. \tag{25}$$

When estimating this specification of the model using likelihood methods, it is convenient for the Phillips curve to include an exogenous stochastic shock. While the microfoundations do not support a shock on the Phillips curve when expressed in terms of the output gap, I amend equation (25) with a "cost push" shock, so the Phillips curve that is estimated has the following form,

$$\pi_t = \left(\frac{1}{1-\beta\gamma}\right)\left[\gamma\pi_{t-1} + \beta E_t \pi_{t+1} + \kappa\left(\mu \tilde{y}_t - \tilde{\lambda}_t\right) + u_t\right], \tag{26}$$

where u_t is an exogenous cost push shock that evolves according to,

$$u_t = \rho_u u_{t-1} + \epsilon_{u,t} \tag{27}$$

where $\epsilon_{u,t}$ is independently and identically distributed with mean zero and variance given by σ_u^2.

2.2.4 Monetary Policy

The nominal interest rate is determined jointly with output and inflation by monetary policy. In this paper I assume the monetary authority follows a Taylor (1993) type rule where the interest rate is set in response to output and inflation, with a preference for interest rate smoothing. In the specification of the model without capital accumulation, the monetary authority is assumed to respond to the output gap, therefore the Taylor rule is given by,

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi \pi_t + \psi_y \tilde{y}_t\right) + \epsilon_{r,t} \tag{28}$$

where $\rho_r \in [0,1)$ is a degree of interest rate smoothing desired by the monetary authority, $\psi_\pi \in (0,\infty)$ is the feedback on the interest rate to inflation, $\psi_y \in (0,\infty)$ is the feedback on the interest rate to output gap, and $\epsilon_{r,t}$ is an independently and identically distributed exogenous monetary policy shock with mean zero and variance given by σ_r^2.

In the model specification with endogenous capital accumulation, the monetary authority is instead assumed to respond to the deviation of output from the steady state (instead of the deviation from the flexible price outcome). In this case the Taylor rule is given by,

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi \pi_t + \psi_y \hat{y}_t\right) + \epsilon_{r,t}. \tag{29}$$

2.3 Complete Model

The complete system for the specification without capital accumulation has four variables: the output gap ($\tilde{y}_t$), the marginal utility of income gap ($\tilde{\lambda}_t$), the inflation rate (π_t), and the interest rate (r_t). The marginal utility is defined by equation (4), and the output gap, inflation rate, and interest rate, are jointly determined by the Euler equation (5), the Phillips curve (26), and the Taylor rule (28). The stochastic shocks in the system are the natural interest rate shock, whose evolution is given in equation (6), the cost push shock whose given in equation (27), and the monetary policy shock.

The model specification that includes endogenous capital accumulation has eight variables: consumption ($\hat{c}_t$), marginal utility of income ($\hat{\lambda}_t$), investment ($\hat{I}_t$), capital stock ($\hat{k}_t$), marginal cost ($\hat{s}_t$), output ($\hat{y}_t$), inflation (π_t), and the interest rate ($\hat{r}_t$). The demand side of the model consists of the Euler equation, (1), and the definition of the marginal utility of income, (2). The supply side of the model consists of the Phillips curve, (24), the definition of the marginal cost, (15), the evolution of capital, (20), and the relationship between investment and capital, (18). The model is completed with the monetary policy rule, (29), and the following log-linear goods market clearing condition,

$$\hat{y}_t = c_y \hat{c}_t + \delta k_y \hat{I}_t, \tag{30}$$

where c_y is the steady state consumption to output ratio and k_y is the steady state capital to output ratio. Appendix A shows that k_y and c_y are given by,

$$k_y = \frac{\beta\alpha(\theta-1)}{\theta\left(1-\beta+\beta\delta\right)},$$

$$c_y = 1 - \delta k_y.$$

There are four exogenous shocks in the model: the preference shock, ξ_t, whose evolution is given in equation (3); the technology shock, z_t, whose evolution is given in equation (16); the investment shock, μ_t, whose evolution is given in equation (21); and the monetary policy shock, $\epsilon_{r,t}$.

3 Learning

3.1 Learning process

The specific type of adaptive learning process considered in this paper is least squares learning. Under least squares learning, agents form expectations by collecting past data and computing least squares estimates. The specific type of least squares learning I use is constant gain learning, which is consistent with agents' forecasts based on weighted least squares, where more recent observations are given more weight, and the weights decline geometrically with the age of the observations. This is a popular assumption in the learning literature and is the same type of learning used by Orphanides and Williams (2005b) to explain inflation scares, Primiceri (2006) to explain the inflation volatility in the 1970s, and Milani (2005) to explain output and inflation persistence.

Constant gain least squares learning is arguably similar to how expectations are actually formed in the U.S economy. Under some of the specifications of least squares learning examined in this paper, agents in the model use exactly the same data that is available to an econometrician, such as the output gap, the inflation rate, and the federal funds rate. Least squares forecasts also often out-perform more complex economic models in out of sample forecasts, and the welfare of individuals who make output, consumption, and savings decisions depend on the accuracy of forecasts, and not the ability to identify parameters of an econometric model, or the ability to make counter-factual predictions. These latter qualities, found in structural economic models, are desirable mostly by policy makers. The constant gain assumption can also be argued as realistic as it captures the idea that agents agents believe changes in the economy are possible, so that agents view more recent data as more likely to yield accurate forecasts than data from further in the past. I demonstrate in the next section that constant gain least squares is equivalent to a very specific type of weighted least squares which is not an actual popular estimation method. However, Evans and Honkapohja (2001) suggest that constant gain least squares is a good approximation for agents that use a "rolling window" of data. That is agents do not use all the data as far back as possible, but form forecasts based on the most recent data for a given number of observations. This is very close to common practice, as empirical studies that forecast output and inflation typically use at most 50 years of data, despite annual data available from Johnston and Williamson (2007) for both these variables dating all they way back to the year 1790.

There is also a theoretical and empirical appeal to using constant gain learning. The theoretical appeal is that unlike with ordinary least squares, with weighted least squares the effects of learning persist in the long run. With ordinary least squares, as time progresses agents obtain more and more observations and so their sample sizes approach infinity. Therefore, the effect a single new observation has on the agents' estimation results disappears. Constant gain learning instead assumes that a new observation carries the same weight every period, regardless of how much time has progressed. The empirical appeal is that the degree to which learning affects the dynamics of the economy can be determined by estimating a single parameter, the learning gain. Moreover, with appropriate initial conditions for the learning process, constant gain learning nests the rational expectations framework, where rational expectations is the special case where the learning gain is equal to zero. Standard statistical tests that determine if a parameter is significantly different from zero can determine the statistical significance of learning, and formally reject or fail to reject the assumption of rational expectations.

The log-linearized New Keynesian model in the previous section has the following general form:

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi v_t, \tag{31}$$

$$v_t = A v_{t-1} + \epsilon_t \tag{32}$$

where x_t is a vector of minimum state variables (expressed as percentage deviations from their steady state), E_t^* refers to a possibly non-rational expectations operator, v_t is a vector of structural shocks, and ϵ_t is a vector of independently and identically distributed innovations to the shock process. The minimum state variable solution of the model can be written as,

$$x_t = G x_{t-1} + H v_t. \tag{33}$$

To agents that learn with a correctly specified model, the actual values in the matrices G and H are unknown, but agents use the form of equation (33) to estimate future values of x_t by least squares. It is assumed that when agents begin period t, time t observations are not yet realized; therefore agents collect observations up through time period $t-1$. From this agents make least squares forecast, then make consumption, production, investment, and pricing decisions based on these expectations. Only after these decisions are implemented, that is at the end of time period t, do time t observations become available. This is both a realistic assumption, if using quarterly macroeconomic data, and a mathematically simplifying assumption. The latest numbers from statistical agencies such as the Bureau of Labor Statistics are almost always at least one quarter old. It is of great mathematical convenience, because the term $E_t^* x_{t+1}$ in equation (31) is then only a function of observations through period $t-1$. Therefore, solving for x_t in terms of past state variables is straightforward. If instead $E_t^* x_{t+1}$ was a function of x_t, non-linear numerical methods would be needed to solve the model as least squares forecasts are non-linear.

To forecast x_{t+1}, agents estimate G and H by least squares using as regressors variables in the vector x_{t-1}, and the shocks included in v_t. Assuming agents have data available on shocks is not very realistic, but this assumption can be dropped. In Section 4 I estimate the models under both cases, so that when comparing the results from the learning and rational expectations models, it will be clear what results derive from the learning process, and what results derive from assuming that agents have a more limited information set.

Agents do not use all the variables in x_t as regressors, only those that correspond to non-zero columns in G. If an entire column in G is equal to zero, this implies that the past observation in the associated element in x_{t-1} does not influence x_t in the MSV solution. I assume agents know the structural form of the economy and therefore use as explanatory variables only the variables that have non-zero coefficients in G. In the New Keynesian model in the previous section with firm-specific capital, the variables with non-zero coefficients in G include consumption, capital, the inflation rate, and the interest rate. The variables in the model that are not used as explanatory variables are output, marginal cost, marginal utility of income, and investment. These variables do not directly lead to any persistence in the model and are simply non-stochastic linear functions of the other variables that have non-zero coefficients in G, and so are perfectly co-linear with the variables that agents do include in the regression. In the specification without endogenous capital, the variables used for explanatory variables are the output gap, the inflation rate, and the interest rate.

I assume agents also use a constant term in their least squares forecasts. The structural form of the model, (31), does not include a constant, but since this equation is written in terms of percentage deviations from the steady state, including a constant in agents' estimation equations implies agents do not know the steady state values of the economy.

Let Φ_t denote the time t estimate of the all the coefficients to be estimated in the learning process. These coefficients include a vector of constants, the non-zero columns in G, and all the columns in H in the case where shocks are used as explanatory variables. Let Y_t denote the time t dependent variables used in the learning process. Since time t data is not available to agents, $Y_t = x_{t-1}$. Let X_t denote the vector of time t explanatory variables. If agents include the stochastic shocks in their explanatory variables, $X_t' = [1 \; x_{t-2}' v_{t-1}']$, otherwise $X_t' = [1 \; x_{t-2}']$. If agents estimate equation (33) by ordinary least squares, they form the estimate,

$$\Phi_t' = \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau X_\tau' \right)^{-1} \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau Y_\tau' \right). \tag{34}$$

The ordinary least squares estimate Φ_t can be rewritten into the convenient recursive form:

$$\Phi_t = \Phi_{t-1} + g_t (Y_t - \Phi_{t-1} X_t) X_t' R_t^{-1}, \tag{35}$$

$$R_t = R_{t-1} + g_t (X_t X_t' - R_{t-1}), \tag{36}$$

where $g_t = 1/(t-1)$ is the learning gain.[3] The recursive form demonstrates precisely how expectations are adaptive. Agents take the previous period's estimates, Φ_{t-1} and R_{t-1}, and correct them according to the residual between the previous period's forecast and the new

[3]To show this, let $R_t = \frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau X_\tau'$ and $\Phi' = R_t^{-1} \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau Y_\tau' \right)$

observation. The amount of the correction depends on the learning gain. With ordinary least squares and infinite memory, the learning gain approaches zero as time approaches infinity, so the effect new observations have on updating the beliefs of Φ and R diminish as the number of observations already in the sample approaches infinity. Constant gain learning instead assumes that the learning gain g_t remains constant over time. This allows new observations to influence estimation results by the same weight throughout time. If the constant gain is equal to zero, the estimate Φ_t remains at its initial value throughout time. Given an initial value equal to the rational expectations solution, a zero constant learning gain implies rational expectations.

Let $\hat{g}_{0,t}$ denote the estimated constant term in Φ_t, and let $\hat{G}_t$ and $\hat{H}_t$ denote the time t estimate of G and H, respectively, obtained from Φ_t. Agents' expectation of x_{t+1} is given by,

$$E_t^* x_{t+1} = \hat{g}_{0,t} + \hat{G}_t E_t^* x_t + \hat{H} E_t v_{t+1} \tag{37}$$

Note that equation (37) assumes that expectations about future shocks, v_{t+1}, are rational. This is a common simplifying assumption made in learning models. It is possible to allow agents to also estimate the coefficients in the shock process, but the dynamics deriving from this additional complication are negligible. Since time t observations are not yet available to agents, agents must also estimate x_t by least squares. The time t estimate of x_t is given by,

$$E_t x_t^* = \hat{g}_{0,t} + \hat{G}_t x_{t-1} + \hat{H} v_t. \tag{38}$$

Plugging this into equation (37) yields,

$$E_t^* x_{t+1} = (I + \hat{G}_t)\hat{g}_{0,t} + \hat{G}_t^2 x_{t-1} + \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right) v_t. \tag{39}$$

Plugging the agents' forecast, (39), into the structural form of the model, (31), leads to the following actual law of motion for x_t:

$$x_t = \Omega_0^{-1}\Omega_2 \left(I + \hat{G}_t\right) \hat{g}_{0,t} + \Omega_0^{-1} \left(\Omega_1 + \Omega_2 \hat{G}_t^2\right) x_{t-1} + \Omega_0^{-1} \left[\Psi + \Omega_2 \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right)\right] v_t. \tag{40}$$

4 Estimation

4.1 Data

In this section I estimate two specifications of the New Keynesian model. The first is the specification that does not include capital, and the second is the specification does account for endogenous capital accumulation. In the specification without capital, the model is written in terms of the output gap, the inflation rate, and the interest rate. The output gap is measured by the percentage deviation of real GDP from the Congressional Budget Office measure of potential GDP. The inflation rate measured by the annualized percentage change in GDP deflater, and the interest rate measured by the annualized federal funds rate.

For the model with firm-specific capital, the data includes four variables: real consumption expenditures, real gross private domestic investment, and the same measures of inflation and the interest rate. Output in this model is defined as the sum of consumption and investment.

The model with firm-specific investment is written in terms of percentage deviations of consumption and investment from their respective steady states. Since consumption and investment are non-stationary series, assuming there are steady state values of these variables is not valid. Consumption and investment are instead de-trended by removing a trend growth rate. To illustrate with consumption, let g_c be the average growth rate of consumption, let c_t denote de-trended consumption, and let C_t denote the original consumption series. De-trended consumption is determined according to,

$$c_t = \frac{C_t}{(1+g_c)^t}.$$

Investment is likewise de-trended. Quarterly data from the Federal Reserve Bank of St. Louis FRED database is used for the first quarter of 1960 through the last quarter of 2006. To determine initial conditions for the learning process, pre-sample quarterly data is used on these same variables from the first quarter of 1954 through the last quarter of 1959.

4.2 Initial conditions

For both specifications of the New Keynesian model I estimate the model under five different cases for how expectations are formed. Case 1 is rational expectations, and the remaining cases are learning with different assumptions for initial conditions and what variables agents use as explanatory variables in their least squares estimates. Case 2 can be viewed as the closest to rational expectations. Agents learn according to constant gain least squares, but the initial values for the learning matrices Φ and R are equal to the rational expectations solution. Furthermore, in case 2 agents include the structural shocks as explanatory variables. Since agents have the same information set as in case 1, and the initial conditions under both cases are the same, when the constant learning gain is equal to zero, case 2 is equivalent to case 1.

Case 3 makes another incremental step away from rational expectations. Agents again learn according to constant gain least squares, and their initial conditions for the learning matrices are equal to the rational expectations values, but agents are not able to collect data on past shocks in order to use them as explanatory variables.

The final two cases assume the agents have the same information set as case 3, but the initial conditions for the learning process are different from the rational expectations solution. In case 4 the initial conditions are equal to constant gain least squares estimates from pre-sample data. This is similar to how Milani (2005) initializes the learning matrices, but he uses estimates from a first order vector autoregression estimated from ordinary least squares, which is consistent instead with a decreasing learning gain. In this paper, the initial conditions for the learning process are consistent with the constant learning gain which is estimated jointly with the other parameters of the model.

Equations (35) and (36) describe the least squares learning process with any given learning gain, g_t. When the learning gain is constant, repeated substitution of these equations can show that the coefficient matrix is given by,

$$\Phi_t = \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \right)^{-1} \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} Y'_{t-\tau} \right) \tag{41}$$

In the model specification without capital, $Y'_t = [\tilde{y}_{t-1}\ \pi_{t-1}\ r_{t-1}]$, and $X'_t = [1\ \tilde{y}_{t-2}\ \pi_{t-2}\ r_{t-2}]$. This data agents use in their estimation procedure is observable to the econometrician and in fact is exactly the same data that is used in this paper. To generate an initial condition for Φ_0 for case 4, I evaluate equation (41) using pre-sample data on the output gap, the inflation rate, and the federal funds rate from 1954:Q1 through 1959:Q4.

In the model with firm-specific capital, $Y'_t = [\hat{c}_{t-1}\ \hat{k}_{t-1}\ \pi_{t-1}\ r_{t-1}]$, and $X'_t = [1\ \hat{c}_{t-2}\ \ \hat{k}_{t-2}\ \pi_{t-2}\ \hat{r}_{t-2}]$. In this specification of the model, some of the data agents use are not directly observed by the econometrician. Consumption and the interest rate are expressed as percentage deviations from the steady state, and capital stock is not directly observable. For a given estimate of the consumption to output ratio and the steady state level of output, pre-sample data on aggregate consumption can be expressed in terms of the percentage deviation from the steady state. The steady state level of the interest rate is equal to $(1/\beta - 1)$, so for a given value of β, the federal funds rate can be expressed in terms of the percentage deviation from the steady state. The inflation rate in agents' regressions is observable and is equal to the growth rate of the GDP deflater.

Data for the U.S. capital stock does not exist, but using the New Keynesian model, data for percentage deviation of capital from its steady state level can be composed from data on the deviation of investment from its steady state level. By suppressing the investment shock in the New Keynesian model during the pre-sample, equation (18) implies,

$$\hat{k}_{t+1} = (1-\delta)\hat{k}_t + \delta \hat{I}_t. \tag{42}$$

For a given initial value for $\hat{k}_t$ in the pre-sample, and given pre-sample data on the percentage deviation of investment from the steady state, $\hat{I}_t$, a pre-sample series of $\hat{k}_t$ can be constructed. I suppose at the beginning of the pre-sample capital is equal to its steady state value so that the pre-sample initial value for $\hat{k}_t$ is equal to zero. The steady state value of investment is equal to the investment to output ratio times the steady state level of output. Given estimates of these parameters, gross private domestic investment is converted to percentage deviations from the steady state level. The initial condition for Φ_0 is computed with equation (41) using this constructed pre-sample data for consumption, capital, inflation, and the interest rate.

Finally, case 5 estimates the initial conditions for the elements of the learning matrices jointly with the other parameters of the model. To keep the likelihood maximization procedure tractable, only the initial conditions for the constant terms and the non-zero elements in G are estimated. The initial condition for R remains given by the rational expectations solution. This still introduces many additional parameters to estimate. In the specification without capital, the output gap, inflation, and interest rate are all estimated using a constant and one period past values of these variables. Therefore there are four coefficients for each variable, for a total of twelve additional parameters. In the specification with firm-specific capital, consumption, capital, the inflation rate, and the interest rate all depend on a constant and one period past values of these variables. This comes to five coefficients for each variable, for a total of twenty additional parameters. As the dataset for estimating quarterly macroeconomic models is often relatively small, identifying these initial conditions jointly with almost twenty parameters of the New Keynesian model with reasonable standard errors is impossible. The motivation for proceeding with this estimation is to examine how varying the initial conditions can influence the results for the other parameters in the model.

4.3 Maximum Likelihood Procedure

I estimate the model by maximum likelihood following the Kalman filter procedure outlined in chapter 13 of Hamilton (1994). This procedure involves rewriting the model into state space form. The state equation is a linear equation describing the entire New Keynesian model including the learning mechanism. The equations governing the state are the actual law of motion for x_t, given in equation (40), and the evolution of the structural shocks given in equation (32). Equation (40) can be rewritten more compactly as,

$$x_t = b_t + F_t x_{t-1} + M_t v_t, \tag{43}$$

where vector b_t and matrices F_t and M_t are given by,

$$b_t = \Omega_0^{-1}\Omega_2\left(I + \hat{G}_t\right)\hat{g}_{0,t},$$

$$F_t = \Omega_0^{-1}\left(\Omega_1 + \Omega_2\hat{G}_t^2\right),$$

$$M_t = \Omega_0^{-1}\left[\Psi + \Omega_2\left(\hat{G}_t\hat{H}_t + \hat{H}_t A\right)\right]$$

This equation can be combined with equation (32) into the single state equation,

$$x_t^* = b_t^* + F_t^* x_{t-1}^* + \epsilon_t^*, \tag{44}$$

where $x_t^* = [x_t' \; v_t']'$ and,

$$F_t^* = \begin{bmatrix} F_t & M_t A \\ 0 & A \end{bmatrix},$$

$$b_t^* = \begin{bmatrix} b_t \\ 0 \end{bmatrix},$$

$$\epsilon_t^* = \begin{bmatrix} M_t \epsilon_t \\ \epsilon_t \end{bmatrix}.$$

The variance of ϵ_t^* is given by,

$$Var(\epsilon_t^*) = \begin{bmatrix} M_t \Sigma M_t' & M_t \Sigma \\ \Sigma M_t' & \Sigma \end{bmatrix},$$

where Σ is a diagonal matrix with the variance of the structural shocks along the diagonal.

In the specification of the New Keynesian model without capital, the observation equations are given by,

$$\begin{aligned} GDP_t^{gap} &= 100\tilde{y}_t \\ INF_t &= \pi^* + 400\pi_t \\ FF_t &= \pi^* + 400\left(r^n + \hat{r}_t\right), \end{aligned} \tag{45}$$

where GDP_t^{gap} is the output gap, INF_t is the annualized quarterly growth rate of the GDP deflater, FF_t is the federal funds rate, $r^n = 1/\beta - 1$ is the steady state real interest rate, and π^* is the steady state inflation rate. The derivation of the New Keynesian model relies on the assumption that there exists a steady state price level which implies the steady state level of inflation is equal to zero. Since this is unlikely to be an accurate assumption, I allow the steady state level of inflation to be possibly greater than zero in the maximum likelihood procedure.

This observation equation and the state equation, (44), are used to form the log-likelihood function which is maximized with respect to the vector of parameters,

$$\Theta_1 = [\eta \ \sigma \ \kappa \ \gamma \ \rho_r \ \psi_y \ \psi_\pi \ \rho_n \ \rho_u \ \sigma_n \ \sigma_u \ \sigma_r \ \pi^* \ g].$$

This set of parameters includes all but two parameters of the model: the discount factor β, and the inverse elasticity of the labor supply, μ. The discount factor is calibrated at $\beta = 0.99$. The parameter μ appears in the Phillips curve equation, (26), along with κ multiplying the output gap, and so is not separately identifiable. I assume that labor supply is perfectly elastic, so that $\mu = 0$. This assumption still allows the output gap to influence inflation in the Phillips curve through the goods market clearing channel, but shuts off the influence through the labor market clearing channel.

In the specification with firm specific capital, the observations equations are given by,

$$\begin{aligned} CONS_t &= c_y y^* + c_y y^* \hat{c}_t \\ INV_t &= (1 - c_y)y^* + (1 - c_y)y^* \hat{I}_t \\ INF_t &= \pi^* + 400\pi_t \\ FF_t &= \pi^* + 400\left(r^n + \hat{r}_t\right), \end{aligned} \tag{46}$$

where $CONS_t$ is the de-trended level of consumption, INV_t is the de-trended level of investment, and $y^*$is the steady state levels of output, which is estimated jointly with the other parameters of the model. The log-likelihood function is maximized with respect to the vector of parameters,

$$\Theta_2 = [\eta\ \sigma\ c_y\ \phi\ \kappa\ \gamma\ \rho_r\ \psi_y\ \psi_\pi\ \rho_z\ \rho_\xi\ \rho_\mu\ \sigma_z\ \sigma_\xi\ \sigma_\mu\ \sigma_r\ \pi^*\ y^*\ g].$$

Parameters that are not estimated in include the discount factor, β, the depreciation rate, δ, and the inverse elasticity of labor supply, μ. The discount rate is again calibrated to $\beta = 0.99$. The depreciation rate is calibrated to $\delta = 0.025$, which corresponds to roughly 10% annual depreciation. The inverse elasticity of labor supply is again poorly identified. It appears again in the Phillips curve along with κ multiplying output. It also appears in the evolution of the aggregate capital stock, (20), along with α multiplying future expectations of output and capital. I assume again that labor supply is perfectly inelastic, so that $\mu = 0$. Future expectations of output and capital still influence intermediate goods firms' investment decisions directly, but not indirectly though the labor market clearing channel.

5 Results

5.1 Model Without Capital

I first estimate the specification of the New Keynesian model without capital under the five frameworks for expectations described in the previous section.

Cases 1 and 2:

Table 1 shows the maximum likelihood parameter estimates for rational expectations and for learning where agents include structural shocks as explanatory variables and have the rational expectations solution for the initial conditions. The estimate for the constant learning gain in case 2 is very small ($g = 0.0067$) and statistically insignificant from zero. Recall that rational expectations is the special case of this learning framework where the learning gain is equal to zero. The learning gain not statistically significantly different from zero implies a failure to reject the null hypothesis that agents form expectations rationally.

The other parameters have very similar estimates under case 1 and case 2. There is no evidence that the learning framework in case 2 helps explain macroeconomic persistence. The estimate for the degree of habit formation under rational expectations is 0.4221 and under learning is 0.4241. These estimates are somewhat lower than that found in other empirical studies. Milani (2005) finds under rational expectations an estimate approximately equal to 0.911. Smets and Wouters (2005) estimate a much more complex New Keynesian model under rational expectations for the U.S. and finds somewhat closer esimtates of the degree of habit formation. They find an estimate approximately equal to 0.69 when using quarterly data from 1974 through 2002, and another estimate equal to 0.44 when using quarterly data from 1982 through 2002. Despite relatively low estimates for the degree of habit formation, the estimates are significantly different from zero, indicating the habit formation is a significant source of persistence in the model.

Inflation indexation is another possible source for persistence. The estimates for inflation indexation are very high: approximately equal to 0.99 under both learning and rational expectations. Milani also finds a higher degree of inflation indexation under rational expectations. Other significant sources of persistence include the interest rate smoothing parameter and the persistence in the natural interest rate shock. The cost push shock is not a significant source of either persistence or volatility under either expectations framework. The estimate of the persistence is nearly equal to zero, and the standard deviation of the shock is approximately equal to 0.003, or 0.3%, under learning and rational expectations. This result is reassuring since it is added to the model to make maximum likelihood estimation possible, despite there not being a microfoundation for the shock.

The final rows of Table 1 report the log-likelihood and the mean squared errors (MSE) for the output gap, the inflation rate, and the federal funds rate under rational expectations and learning. Consistent with the small and statistically insiginicant estimate of the learning gain coefficient, the there is only a tiny improvement in the log-likelihood by adding learning. The improvement in the MSE is also negligible.

Case 3:

Table 2 shows the results for learning under case 3, where the initial condition for the learning process is again equal to the rational expectations solution, but agents do not include the structural shocks as explanatory variables; the data they use are the output gap, the inflation rate, and the interest rate. For easy comparison, Table 2 also repeats the estimation results under rational expectations. The estimate for the learning gain is approximately equal to 0.0042, which is again small and statistically insignificant. Since this learning framework does not nest rational expectations, the small learning gain does not imply expectations are rational, only that agents' coefficient estimates are slow to adjust. Since the constant gain learning procedure can be viewed as an approximation to learning with a rolling window of data, Milani (2005), Evans and Honkapohja (2001), and others interpret the inverse of the estimate of the learning gain as the number of observations agents use every period to form their expectations. In this case the learning gain implies that agents use approximately 238 quarterly observations, or 59.5 years of data.

Again, many of the other parameter estimates are similar to the estimates under rational expectations. The most notable exceptions are the inverse of the pseudo intertemporal elasticity of substitution, σ, and the standard deviation of the natural interest rate shock. Under learning, the estimate for σ is 0.2251, compared to 0.5152 under rational expectations. This indicates that the learning model predicts consumption decisions are more sensative to changes in the real interest rate. This is likely due to the fact that consumption decisions are less likely to change in response to stochastic shocks, since in case 3 stochastic shocks are not used to form expectations. Therefore to capture the same degree of consumption (and therefore output) volatility, the learning model predicts a higher intertemporal elasticity of substitution.

A similar story explains the difference in the estimate for the natural interest rate shock. The estimate under the learning model is 0.2736 compared to 0.0751 under the rational expectations model. Since under this learning framework structural shocks do not contemporaneously influence expectations, the size of the natural interest rate shock is larger to capture the same degree of volatility as under rational expectations.

The final rows of Table 2 show that the learning model does not signicantly explain the data better than rational expectations. The log-likelihood is only a tiny bit larger under learning and the MSE of the output gap, inflation rate, and federal funds rate are almost identical under rational expectations and learning.

Case 4:
The results for rational expectations and case 4 are presented in Table 3. In this learning specification, agents again do not use stuctural shocks as explanatory variables and now the initial conditions are equal to weighted least squares estimates from pre-sample data. The estimate of the learning gain is equal to 0.0828, much larger than the previous estimates. This learning gain implies that agents use only 12 observations in forming expectations, or only three years of data. With such a large learning gain agents' coefficient estimates evolve much more quickly. The initial values for agents' coefficients are farther away from the rational expectations solution than under the previous cases, so the larger learning gain allows the coefficients to adjust more rapidly towards the rational expectations solution. The final rows of Table 3 indicate that imposing these initial conditions worsen the performance of the learning model. The log-likelihood is lower under learning than under rational expectations, and the MSE of inflation and the federal funds rate are higher under learning.

The most notable differences in the parameter estimates between learning and rational expectations include the pseudo intertemporal elasticity of substitution, the standard deviation of the natural interest rate shock, and the monetary policy parameters on inflation and the output gap. Again the learning model predicts a higher intertemporal elasticity of substitution and a more volatile natural rate shock to compensate structural shocks not contemporaneously impacting expectations. The monetary policy parameters for the feedback on the output gap and inflation are lower in the learning framework. This indicates the initial conditions generate more movement in output and pricing decisions. The monetary authority's reaction to these variables is predicted to be smaller to match the volatility of the federal funds rate.

Case 5:
Table 4 reports the estimation results for the parameters when the initial conditions of the learning process are estimated jointly with all the other parameters. The estimated values of the initial conditions are not reported as they have very large standard errors and there is no clear interpretation of the coefficients from the reduced form model agents use to compute their forecasts. The most notable differences in the parameter estimates is not the point estimates, but the standard deviations of the parameter estimates. The standard deviations of the estimates for the degree of habit formation, the inverse of the pseudo intertemporal elasticity of substitution, the degree of price indexation, and the standard deviation of the natural interest rate shock are all very large. This indicates that the point estimates of these parameters can completely depend on how initial conditions for the learning process

are specified. Moreover the sizable standard deviations on these parameter estimates combined with the sizable standard deviations on estimated initial conditions imply that jointly identifying the parameters and the initial conditions is impossible.

Comparisons:

The results from these five expectations frameworks show that learning does not explain U.S. data any better than rational expectations, at least in terms of the simple New Keynesian model. Figure 1 shows the paths of the forecast errors for the output gap, the inflation rate, and the federal funds rate under each expectations framework over the sample period to determine if learning may better explain different periods in post-war U.S. history. The top row shows the rational expectations model makes the largest forecast errors during the 1970s decade for the output gap and the inflation rate. The largest forecast errors for the federal funds rate are made from 1979 through 1982, the begining of Paul Volker's tenure as Federal Reserve chairman. Lubik and Schorfheide (2004) find empirical evidence that it was during this period the Federal Reserve policy changed to more aggressively respond to the inflation rate.

The second row of graphs indicate that the learning model under case 2 makes the same forecast errors as the rational expectations model. The number in parantheses is the correlation of the forecast error under learning to the forecast error under rational expectations. The correlations for the forecast errors in case 2 are all above 0.99 indicating the forecast errors under this learning framework are virtually the same as rational expectations. The same is true for learning in case 3 where agents do not use shocks as explanatory variables in their regressions. Despite there being some minor differences in the parameter estimates described above, the correlations indicate the forecast errors are virtually the same as the rational expectations model. The fourth row of graphs show the forecast errors for the output gap and the inflation rate under learning with pre-sample initial conditions are less correlated with the rational expectations forecast errors, but this is because this learning model actually performed worse than rational expectations. The largest output gap and inflation forecast errors are again made during the 1970s, suggesting that learning is no better than rational expectations at explaining the experience of excessively volatile and persistent inflation during this period. The final row of forecast errors show that increasing the number of free parameters to include all the initial conditions for the learning process does not significantly improve the fit of the model. The correlations are still very close to one, and the largest errors occur during the 1970s decade.

Figure 2 shows the series of shocks predicted by each of the expectations framework, along with the correlations with the rational expectations counterpart given in parentheses. The evolution of shocks is virtually identical for rational expectations and learning under cases 2 and 3. The path of the shocks indicate that the volatile inflation during the 1970s is explained by an excessively volatile cost push shock and rather large increases in the natural real interest rate during the early 1970s and middle to late 1970s.

The paths of the shocks with pre-sample initial conditions are not as correlated with rational expectations as the first two learning models. This learning framework still makes the same qualitative prediction that the 1970s was characterized by an excessively volatile cost push shock. The final row of graphs indicate that even estimating the initial conditions jointly with the other parameters does not alter the predicted paths of the shocks.

All the results of this subsection indicate that learning does not improve the performance of the simple New Keynesian model, nor significantly alter the model's predictions. Figure 3 sheds some light as to why that is. This figure shows the evolution of agents expectations under rational expectations and each learning framework. In the second through the fifth rows the solid lines are the expectations predicted under learning and the dotted lines are the expectations predicted by rational expectations. The graphs show that the expectations under learning are very nearly equal to the rational expectations. Since the learning frameworks do not significantly alter expectations, learning is no better able to explain the data than rational expectations, at least for the standard New Keynesian model.

5.2 Model with Capital

I now turn to estimating the New Keynesian model with endogenous capital accumulation to determine how sensative the results are to the model specification. Including capital may increase the role learning can play in the economy, as consumption and output expectations play seperate roles in the economy. Also, an additional explanatory variable, capital, becomes available for agents to use for forecasting. Introducing capital may also lessen the empirical significance of learning as it provides an additional source of output persistence.

Cases 1 and 2:

Table 5 shows the parameter estimates for the model under rational expectations and under learning, case 2. There are some notable differences from the estimation results of the more simple New Keynesian model. The estimate for the degree of habit formation is equal to 0.9181 which indicates habit formation is an even more significant source of persistence. The degree of inflation indexation is very low, equal to 0.0001. Inflation persistence instead is explained by the dependence of the marginal cost on the current period capital stock, as evident from equation (15). The current period capital stock in turn depends on last period's capital stock and investment decisions.

The estimates for the monetary policy parameters are much lower under this specification of the model. The estimate of the responses to inflation and output is 1.0014 and 0.1003, respectively. This difference is at least partially due to using different data in the estimation procedure for output. In the specification of the New Keynesian model without capital, monetary policy responds to the output gap as measured by the deviation of the CBO measure of potential GDP. In the specification with capital, monetary policy responds to the percentage deviation of ouput from its steady state. Steady state output is estimated with y^*, and output is defined as the sum of consumption and investment.

The estimate for the steady state level of inflation is also much lower than in the previous specification. The estimate is equal to 0.2209 and is not significantly different from zero. This suggests that introducing investment dynamics helps explain large increases in inflation present in the data without having to assume a high steady state level of inflation.

Table 5 shows the parameter estimates for rational expectations and learning under case 2 are exactly the same. The estimated constant learning gain is equal to 0.0000, which is the special case for rational expectations, indicating that allowing for such a small deviation from rational expectations does nothing to further explain the data.

Case 3:

Learning under case 3 takes a more substantial step away from rational expectations as agents have a limited information set. Table 6 shows the parameter estimated under rational expectations and case 3 learning. The learning gain is rather small, 0.0052, but it is significantly different from zero indicating that agents expectations do evolve. This estimate coincides with agents using approximately quarterly 192 observations, or 48 years of data, which is rather close to most empirical work using quarterly aggregate data.

Parameter estimates for the standard deviation of the technology and investment shocks are larger under this learning model than under rational expectations. Like in the model without capital, when agents no longer use the shocks as explanatory variables, shocks do not concurrently affect expectations, so the standard deviation of these shocks have larger estimates to generate the same volatility.

The only other difference is the estimate for the steady state level of inflation is higher in case 3. The estimate is equal to 3.8982 and is significantly different from zero. Since shocks do not concurrently influence expectations, it will be seen later that this leads to very less volatile expectations, especially during the 1970s. Therefore, for the maximum likelihood procedure finds a larger steady state level of inflation to better account for this time period.

The final rows in Table 6 show that moving from rational expectations to case 3 learning leads to an improvement in the likelihood, and an improved fit for aggregate investment, but a worse fit for consumption. The MSE's for inflation and the federal funds rate are nearly the same.

Case 4:

Table 7 shows the results for rational expectations and learning with initial conditions set to weighted least squares estimates from pre-sample data. The estimate of the learning gain is 0.0060, which is much smaller than the case 4 estimate under the New Keynesian model without capital. This implies that agents coefficient estimates do evolve, but much more slowly than predicted before. This suggests that the weighted least squares estimates obtained from pre-sample data are more appropriate for the model with capital.

Again the estimates of the standard deviations of the technology and investment shocks are larger than under rational expectations. The estimate for the steady state level of inflation is very close to zero, indicating the initial conditions help explain the run-up of inflation during the 1970s without depending on a higher normal level of inflation. The likelihood and MSE for consumption, investment, and inflation indicate that fixing the initial conditions based on pre-sample data provides a worse fit to the data than the rational expectations model.

Case 5:

Table 8 shows the estimation results for the final case, where the initial conditions are estimated simultaneously with other parameters of the model. Like in the specification without capital, the point estimates for the coefficients under rational expectations and learning are very similar. The standard deviations for the parameters under learning are much smaller than case 5 in model specification without capital, indicating jointly identifying the initial conditions is much more feasible when using a richer model with more data. The standard deviations for estimates for the steady state level of inflation, the degree of habit formation, the inverse of the pseudo intertemporal elasticity of substitution, and the degree of inflation indexation are still rather large, though not as large before. This re-enforces the claim that imposing ad-hock initial conditions for the learning process can drive the point estimates of these parameters. The smaller magnitude of the standard deviations suggest though that joint identification may be possible by further extending the model and incorporating more data.

Comparisons:

Figure 4 shows all the series of forecast errors to determine if any of the learning models with capital are any better at explaining specific periods of U.S. history. The numbers in parentheses are the correlations of the series with the rational expectations counterpart. Rational expectations and learning in case 1 indicate the largest forecast errors for inflation are made during the 1970s. The forecast errors for consumption and investment do not appear clustered around a specific episode. The correlations of the forecast errors for case 3 indicate nearly the same predictions as rational expectations, except for investment. Recall the MSE for investment is smaller under case 3 than rational expectations.

The fourth row shows the forecast errors for the model with pre-sample initial conditions. Notice that the largest forecast errors appear at the begining of the sample. Recall the MSE for consumption, investment, and inflation are larger under learning with pre-sample initial conditions than under rational expectations. The largest errors occur at the begining of the sample, when expectations are most dependent on the initial conditions. As time progresses, the effect of the initial conditions diminish, and the graphs show the forecast errors therefore improve. The final row shows the forecast errors when the initial conditions are estimated. Table 8 shows MSE's are smaller in this case, but the path of the forecast errors are still strongly correlated with the rational expectations forecast errors.

Figure 5 shows the evolution of the structural shocks under each expectations framework. For rational expectations, the predicted path for the technology shock is negative throughout most of the sample, hitting the biggest lows in the early to middle 1970s and again in the early 1980s. On the other hand, the investment shock is predominantly positive until the recessions in 1991 and 2001. The predictions under case 3 learning is very different. The correlation in parentheses in Figure 5 indicates that the investment shock is actually strongly negatively correlated with the path under rational expectations. The recessions in 1991 and 2001 instead coincide with positive investment shocks. The same is true under the remaining learning cases, which all use the same information set as case 3.

Figure 6 sheds some light as to why this happens. The figure shows the evolution of expectations of consumption, inflation, capital stock, and output under each learning framework. The second through fifth rows show the learning expectations (solid line) and the rational expectations (dotted line). The third row of the figure shows the expectations under case 3 for consumption and output are somewhat similar under learning and rational expectations, but the expectation for capital stock is much different. There is very little volatility in the expectations and the expectations are below the rational expectations for much of the sample. The recessions in 1991 and 2001 coincide with the low points in capital stock expectations. Equation (20) shows that capital stock expectations play a role in investment decisions when there is a positive capital adjustment cost. At these low points in capital stock expectations, firms are less inclined to invest in new capital stock as they expect the capital adjustment costs will be higher. The decrease in the demand for investment that results is the explanation for these resessions for case 3, instead of a negative stochastic investment shock.

Notice also from figure 6 that inflation expectations in case 3 are also very low and not very volatile throughout the sample, since stochastic shocks do not influence expectations. To generate the high levels of inflation during the 1970s, case 3 predicts a higher estimate for the steady state inflation rate as noted above. The remaining expectations that are different from rational expectations are difficult to intrepret, but it is worthwhile to note that the New Keynesian model with capital does result is some very different paths for expectations for some variables, which is not true in the model without capital. As a consequence, incorporating learning into the model is able to produce some different explanations for what caused various downturns in the U.S. business cycle, even though learning does not substantially improve the fit of the models to the data.

6 Conclusion

This paper examines a standard New Keynesian monetary model and an extended model that accounts for endogenous firm-specific capital accumulation under rational expectations and four assumptions for constant gain learning. The four frameworks for constant gain learning differ based on the initial conditions for the learning process and the information set available to agents to form expectations. The models are estimated by maximum likelihood and the results indicate that learning provides minimal to no improvement in the fit of the New Keynesian models to U.S. data. When the learning procedure is initialized using least squares esimtation results from pre-sample data, the learning model actually fits the data worse than rational expectations. When assuming that agents have a limited information set, where data on the stochastic shocks is not available to agents, the learning model leads to small improvements in the fit to the data. This assumption also causes some expectations to not be as volatile which leads to some different predictions for the size of the structural shocks hitting the economy, and in the model with endogenous capital, this leads to different explanations for what caused recent recessions in U.S. history. The initial conditions for the learning process are finally estimated jointly with the other parameters of the model. These results show that the initial conditions and some key parameters are not jointly identifiable, and it may be possible that imposing ad-hock initial values can drive the estimates of some of the parameters.

Plots of the forecast errors, estimated structural shocks, and agents expectations are examined to determine if the learning models out perform rational expectations for certain periods of U.S. history. The forecast errors for all the models are largest during the 1970s episode of volatile and persistent output and inflation. Furthermore, the forecast errors of the learning models are highly correlated with the forecast errors from the ratinal expectations model, suggesting that least squares learning in the context of the New Keynesian model does not explain any periods of U.S. history better than rational expectations.

Appendix A Derivation of New Keynesian model

A.1 Consumers

The first order conditions for c_t, n_t, and b_t are

$$\lambda_t = \xi_t \left(c_t - \eta c_{t-1}\right)^{-\sigma} - \beta\eta E_t \xi_{t+1} \left(c_{t+1} - \eta c_t\right)^{-\sigma}$$

$$n_t(i)^{\mu} = \lambda_t \frac{w_t(i)}{p_t}$$

$$\lambda_t = \beta E_t \lambda_{t+1} \frac{1 + r_t}{1 + \pi_{t+1}}$$

where λ_t is the Lagrange multiplier for the budget constraint and therefore the marginal utility of real income. Log-linearizing the first order conditions yields,

$$\hat{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)} \left[\beta\eta\sigma E_t \hat{c}_{t+1} - \sigma(1+\beta\eta^2)\hat{c}_t + \sigma\eta\hat{c}_{t-1}\right] + \left(\hat{\xi}_t - \beta\eta E_t \hat{\xi}_{t+1}\right) \tag{A1}$$

$$\hat{w}_t(i) - \hat{p}_t = \mu \hat{n}_t(i) - \hat{\lambda}_t \tag{A2}$$

$$\hat{\lambda}_t = E_t \hat{\lambda}_{t+1} + \hat{r}_t - E_t \pi_{t+1} \tag{A3}$$

where a hat indicates the percentage deviation of the variable from its steady state. Equation (A2) will be referenced later to express equilibrium real wages in terms of employment. Equations (A3) and (A1) together implicitly define the log-linear Euler equation which determines consumers' demand for final goods.

When investment is not included in the model, the percentage deviation of consumption from its steady state is exactly equal to the percentage deviation of output from its steady state, and therefore the demand side of the economy can be expressed in terms of the output gap, the percentage deviation of output from its flexible price outcome. Let $\hat{x}_t^f$ be the outcome for some variable $\hat{x}_t$ under fully flexible prices, and let $\tilde{x}_t \equiv \hat{x}_t - \hat{x}_t^f$ be the percentage deviation of x_t from the fully flexible outcome. Since the derivations in this section do not depend on the flexibility of prices, equations (A1) and (A3) hold under flexible prices. Re-writing these conditions for fully flexible prices produces,

$$\hat{\lambda}_t^f = \frac{1}{(1-\beta\eta)(1-\eta)} \left[\beta\eta\sigma E_t \hat{c}_{t+1}^f - \sigma(1+\beta\eta^2)\hat{c}_t^f + \sigma\eta\hat{c}_{t-1}^f\right] + \left(\hat{\xi}_t - \beta\eta E_t \hat{\xi}_{t+1}\right), \tag{A4}$$

$$\hat{\lambda}_t^f = E_t \hat{\lambda}_{t+1}^f + \hat{r}_t^f - E_t \pi_{t+1}^f. \tag{A5}$$

Notice, $\hat{\xi}_t$ does not have an f subscript since it is an exogenous shock that is not influenced by the flexibility of prices. Subtracting (A4) from (A1) and (A5) from (A3) and imposing goods market clearing produces,

$$\tilde{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\tilde{y}_{t+1} - \sigma(1+\beta\eta^2)\tilde{y}_t + \sigma\eta\tilde{y}_{t-1}\right] \tag{A6}$$

$$\tilde{\lambda}_t = E_t\tilde{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1} - r_t^n \tag{A7}$$

where $r_t^n \equiv \hat{r}_t^f - E_t\pi_{t+1}^f$ is what is known as the natural interest rate, the real interest rate that would occur under fully flexible prices. This variable is usually assumed to follow an exogenous stochastic process, therefore equations (A6) and (A7) are solely in terms of the stationary variables: output gap, inflation, and interest rate.

A.2 Producers

A.2.1 Final goods firms

The final goods firm chooses its demand for intermediate good i to maximize profits,

$$\Pi_t = p_t\left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}}di\right]^{\frac{\theta}{\theta-1}} - \int_0^1 p_t(i)y_t(i)di$$

The first order condition leads to the demand for intermediate good i,

$$y_t(i) = \left[\frac{p_t(i)}{p_t}\right]^{-\theta} y_t. \tag{A8}$$

which is given in equation (8).

A.2.2 Input choices

Intermediate goods firms choose labor demand and rent capital to minimize real total cost, given in equation (14), subject to the production function, given in equation (13). The first order conditions are,

$$\frac{w_t(i)}{p_t} = (1-\alpha)s_t(i)\frac{y_t(i)}{n_t(i)}, \tag{A9}$$

$$\rho_t(i) = \alpha s_t(i)\frac{y_t(i)}{k_t(i)}, \tag{A10}$$

where $s_t(i)$ is the Lagrange multiplier on the production function. The Lagrange multiplier is interpreted as the change in the objective function from a marginal ease in the constraint. In this case the objective function is total cost and the constraint is total output, so the Lagrange multiplier is equal to the marginal cost.

Log-linearizing the first order conditions yields,

$$\hat{\rho}_t(i) = \hat{s}_t(i) + \hat{y}_t(i) - \hat{k}_t(i), \tag{A11}$$

$$\hat{w}_t(i) - \hat{p}_t = \hat{s}_t(i) + \hat{y}_t(i) - \hat{n}_t(i), \tag{A12}$$

Combining these two equations to eliminate $\hat{s}_t(i)$ and substituting equation (A2) to eliminate wages and prices leads to the expression for the rental rate of capital,

$$\hat{\rho}_t(i) = (\mu + 1)\hat{n}_t(i) - \hat{k}_t(i) - \hat{\lambda}_t. \tag{A13}$$

The production function can now be used to express the rental rate of capital only in terms of output and capital. The log-linear production function is given by,

$$\hat{y}_t(i) = \hat{z}_t + \alpha\hat{k}_t(i) + (1-\alpha)\hat{n}_t(i). \tag{A14}$$

Solving equation (A14) for $\hat{n}_t(i)$ and substituting this into (A13) yields,

$$\hat{\rho}_t(i) = \frac{\mu+1}{1-\alpha}\hat{y}_t(i) - \frac{1+\alpha\mu}{1-\alpha}\hat{k}_t(i) - \hat{\lambda}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t \tag{A15}$$

Solving equation (A11) for $\hat{s}_t(i)$ and using equation (A15) to substitute out $\hat{\rho}_t(i)$ leads to the expression for marginal cost for firm i,

$$\hat{s}_t(i) = \frac{\mu+\alpha}{1-\alpha}\hat{y}_t(i) - \frac{\alpha(\mu+1)}{1-\alpha}\hat{k}_t(i) - \hat{\lambda}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t \tag{A16}$$

Summing over all the firms leads to the average marginal cost in the economy,

$$\hat{s}_t = \frac{\mu+\alpha}{1-\alpha}\hat{y}_t - \frac{\alpha(\mu+1)}{1-\alpha}\hat{k}_t - \hat{\lambda}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t \tag{A17}$$

In the specification of the model that does not allow for endogenous capital accumulation $\alpha = 0$, therefore the marginal cost in this case is given by,

$$\hat{s}_t = \mu\hat{y}_t - \hat{\lambda}_t - (\mu+1)\hat{z}_t \tag{A18}$$

Subtracting equation (A17) from equation (A16), leads to an expression for the marginal cost of firm i in terms of the average marginal cost and the firms relative output and capital stock,

$$\hat{s}_t(i) = \hat{s}_t + \frac{\mu+\alpha}{1-\alpha}\left[\hat{y}_t(i) - \hat{y}_t\right] - \frac{\alpha(\mu+1)}{1-\alpha}\tilde{k}_t(i) \tag{A19}$$

where $\tilde{k}_t(i) = \hat{k}_t(i) - \hat{k}_t$ is the relative capital stock of firm i.

A.2.3 Capital goods firms

Capital goods firms maximize the utility value of profits, given in equation (19), subject to the evolution of firm-specific capital stock, given in equation (17). Instead of explicitly computing the profit maximizing choice of investment, one can solve the evolution of capital for $I_t(i)$ and substitute this into the objective function. The first order condition is,

$$\begin{aligned}
&\frac{\lambda_t}{\mu_t}\left[1+\phi\left(\frac{k_{t+1}(i)}{k_t(i)}-1\right)\right]= \\
&\beta E_t\frac{\lambda_{t+1}}{\mu_{t+1}}\left[\mu_{t+1}\rho_{t+1}(i)+(1-\delta)+\phi\left(\frac{k_{t+2}(i)}{k_{t+1}(i)}-1\right)\frac{k_{t+2}(i)}{k_{t+1}(i)}-\frac{\phi}{2}\left(\frac{k_{t+2}(i)}{k_{t+1}(i)}-1\right)^2\right].
\end{aligned} \tag{A20}$$

Log-linearizing this yields,

$$\begin{aligned}
&\hat{\lambda}_t+\phi\left(\hat{k}_{t+1}(i)-\hat{k}_t(i)\right)=E_t\hat{\lambda}_{t+1}+\left[1-\beta\left(1-\delta\right)\right]E_t\hat{\rho}_{t+1}(i) \\
&+\beta\phi\left(E_t\hat{k}_{t+2}(i)-\hat{k}_{t+1}(i)\right)+\hat{\mu}_t-\beta(1-\delta)E_t\hat{\mu}_{t+1}.
\end{aligned} \tag{A21}$$

Plugging equation (A15) into (A21) leads to the following equilibrium condition for the evolution of capital stock for firm i:

$$\begin{aligned}
&\hat{\lambda}_t+\phi\left(\hat{k}_{t+1}(i)-\hat{k}_t(i)\right)=\beta(1-\delta)E_t\hat{\lambda}_{t+1} \\
&+\left(\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right)\left[(\mu+1)E_t\hat{y}_{t+1}(i)-(1+\alpha\mu)\hat{k}_{t+1}(i)\right] \\
&+\beta\phi\left(E_t\hat{k}_{t+2}(i)-\hat{k}_{t+1}(i)\right)-\frac{(\mu+1)\left[1-\beta\left(1-\delta\right)\right]}{1-\alpha}E_t\hat{z}_{t+1}+\hat{\mu}_t-\beta(1-\delta)E_t\hat{\mu}_{t+1}.
\end{aligned} \tag{A22}$$

Integrating equation (A22) over all firms leads to the evolution of the aggregate capital stock,

$$\begin{aligned}
&\hat{\lambda}_t+\phi\left(\hat{k}_{t+1}-\hat{k}_t\right)=\beta(1-\delta)E_t\hat{\lambda}_{t+1} \\
&+\left(\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right)\left[(\mu+1)E_t\hat{y}_{t+1}-(1+\alpha\mu)\hat{k}_{t+1}\right] \\
&+\beta\phi\left(E_t\hat{k}_{t+2}-\hat{k}_{t+1}\right)-\frac{(\mu+1)\left[1-\beta\left(1-\delta\right)\right]}{1-\alpha}E_t\hat{z}_{t+1}+\hat{\mu}_t-\beta(1-\delta)E_t\hat{\mu}_{t+1}.
\end{aligned} \tag{A23}$$

Subtracting equation (A23) from equation (A22) leads to following expression for firm i's capital stock in terms of the aggregate capital stock,

$$\begin{aligned}
&\phi\left(\tilde{k}_{t+1}(i)-\tilde{k}_t(i)\right)=\beta\phi\left(E_t\tilde{k}_{t+2}(i)-\tilde{k}_{t+1}(i)\right) \\
&+\left[\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right]\left[(\mu+1)E_t\left(\hat{y}_{t+1}(i)-\hat{y}_{t+1}\right)-(1+\alpha\mu)\tilde{k}_{t+1}(i)\right].
\end{aligned} \tag{A24}$$

A.2.4 Optimal pricing

The inflation indexation rule given in equation (23) can be re-written so that future prices intermediate goods firms will charge while not being able to re-optimize their price can be expressed in terms of the price chosen by the firm at time t. By repeated substitution of equation (23), the price at time $t+T$ of good i can be expressed as,

$$p_{t+T}(i) = p_t(i)\exp\left(\gamma\sum_{\tau=0}^{T-1}\pi_{t+\tau}\right),$$

For notational convenience, let $\pi^*_{t+T} \equiv \sum_{\tau=0}^{T-1}\pi_{t+\tau}$. Substitute the demand equation, (8), into the profit function, (22), to express the profit only in terms of the intermediate good price, $p_t(i)$, and aggregate state variables the firm cannot control:

$$E_t\sum_{T=0}^{\infty}(\omega\beta)^T\frac{\lambda_{t+T}}{\lambda_t}\left\{\left(\frac{p_t(i)e^{\gamma\pi^*_{t+T}}}{p_{t+T}}\right)^{1-\theta}y_{t+T} - S\left[\left(\frac{p_t(i)e^{\gamma\pi^*_{t+T}}}{p_{t+T}}\right)^{-\theta}y_{t+T}\right]\right\}. \tag{A25}$$

The first order condition with respect to $p_t(i)$ is given by,

$$E_t\sum_{T=0}^{\infty}(\omega\beta)^T\frac{\lambda_{t+T}}{\lambda_t}\left\{(1-\theta)\left(\frac{p^*_t(i)e^{\gamma\pi^*_{t+T}}}{p_{t+T}}\right)^{1-\theta} + \theta s_{t+T}(i)\left(\frac{p^*_t(i)e^{\gamma\pi^*_{t+T}}}{p_{t+T}}\right)^{-\theta}\right\}\frac{y_{t+T}}{p^*_t(i)} = 0, \tag{A26}$$

where $p^*_t(i)$ is the optimal price for a firm that is able to re-optimize its price. Since the first order condition cannot be rewritten in terms of inflation instead of prices, it is necessary to assume prices have a steady state, which implies the steady state level of inflation is equal to zero. Before log-linearizing, it is convenient to rearrange equation (A26) as,

$$\begin{aligned}(1-\theta)\,E_t\sum_{T=0}^{\infty}(\omega\beta)^T\lambda_{t+T}\left(\frac{p^*_t(i)e^{\gamma\pi^*_{t+T}}}{p_{t+T}}\right)^{1-\theta}y_{t+T} &= \\ -\theta E_t\sum_{T=0}^{\infty}(\omega\beta)^T\lambda_{t+T}s_{t+T}(i)\left(\frac{p^*_t(i)e^{\gamma\pi^*_{t+T}}}{p_{t+T}}\right)^{-\theta}y_{t+T}&,\end{aligned} \tag{A27}$$

then log-linearize each side of the equal sign separately. Log-linearizing the left hand side and right hand size, respectively, yield,

$$(1-\theta)\,\lambda y E_t\sum_{T=0}^{\infty}(\omega\beta)^T\left[\hat{\lambda}_{t+T} + \hat{y}_{t+T} + (1-\theta)\left(\hat{p}^*_t(i) - \hat{p}_{t+T} + \gamma\pi^*_{t+T}\right)\right], \tag{A28}$$

$$-\theta\lambda y s E_t\sum_{T=0}^{\infty}(\omega\beta)^T\left[\hat{\lambda}_{t+T} + \hat{y}_{t+T} + \hat{s}_{t+T} - \theta\left(\hat{p}^*_t(i) - \hat{p}_{t+T} + \gamma\pi^*_{t+T}\right)\right] \tag{A29}$$

where λ is the steady state marginal utility of income, y is the steady state level of output, and s is the steady state marginal cost. Steady state marginal utility and steady output cancel out from the left and right hand sides. The steady state marginal cost is found by evaluating the first order condition (A26) where $\lambda_t = \lambda$ and $p_t^*(i) = p_t = p$ for all t. In the steady state equation (A26) simplifies to,

$$\left(\frac{1}{1-\omega\beta}\right)\frac{(1-\theta+\theta s)\,y}{p} = 0.$$

The steady state solution for s is given by,

$$s = -\frac{1-\theta}{\theta}. \tag{A30}$$

The coefficient $-\theta\ s$ in equation (A29) therefore cancels out with $1-\theta$ in equation (A28). Combining the left and right hand side then yields,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* - \hat{s}_{t+T}(i)\right] = 0 \tag{A31}$$

Solving for $\hat{p}_t^*(i)$ yields,

$$\hat{p}_t^*(i) = (1-\omega\beta)\,E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \hat{s}_{t+T}(i)\right]. \tag{A32}$$

Substitute into equation (A32), the log-linearized the demand for intermediate good i at time $t+T$, which is given by,

$$\hat{y}_{t+T}(i) = -\theta(\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^*) + \hat{y}_{t+T} \tag{A33}$$

and the marginal cost given in equation (A19). This leads to an expression for the optimal price for firm i in terms of aggregate variables and the firm's expected future capital,

$$\begin{aligned}\hat{p}_t^*(i) = \quad & (1-\omega\beta)\,E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left\{\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \hat{s}_{t+T} - \frac{\theta\,(\mu+\alpha)}{1-\alpha}\left[\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^*\right]\right\} \\ & -(1-\omega\beta)\,E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left\{\frac{\alpha\,(\mu+1)}{1-\alpha}\tilde{k}_{t+T}(i)\right\}.\end{aligned} \tag{A34}$$

The solution of this equation for $\hat{p}_t^*(i)$ is given by,

$$\hat{p}_t^*(i) = (1-\omega\beta)\,E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \psi\hat{s}_{t+T} - \frac{\psi\alpha(\mu+1)}{1-\alpha}\tilde{k}_{t+T}(i)\right], \tag{A35}$$

where

$$\psi = \left[1 + \frac{\theta(\mu+\alpha)}{1-\alpha}\right]^{-1}.$$

Equation (A35) can be rewritten as the first order difference equation:

$$\hat{p}_t^*(i) = \omega\beta E_t \hat{p}_{t+1}^*(i) + (1-\omega\beta)\left(\hat{p}_t + \psi\hat{s}_t - \frac{\psi\alpha(\mu+1)}{1-\alpha}\tilde{k}_t(i)\right), \tag{A36}$$

where $E_t\hat{p}_{t+1}^*(i)$ denotes the expectation at time t for the time $t+1$ optimal decision for the firm's new price, conditional that the firm is able to re-optimize its price again in period $t+1$. Note, this is not the same as the unconditional time t expectation of the firm's price in period $t+1$. Since with probability ω the firm will not be able to re-optimize its price next period, the unconditional expectation for firm i's price in period $t+1$ is given by,

$$E_t\hat{p}_{t+1}(i) = \omega\left[\hat{p}_t^*(i) + \gamma\pi_{t-1}\right] + (1-\omega)E_t\hat{p}_{t+1}^*(i). \tag{A37}$$

A.2.5 Phillips Curve Solution Without Capital

Equation (A36) expresses the solution of the optimal price for firm i in terms of aggregate variables $\hat{p}_t$ and $\hat{s}_t$ and the relative capital stock of firm i, $\tilde{k}_t(i)$. When α, the exponent on capital in the production function is equal to zero, the coefficient on the relative capital stock in the optimal price equation is equal to zero. In this special case, equations (A35) and (A36) simplify to, respectively,

$$\hat{p}_t^*(i) = (1-\omega\beta)\, E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \frac{1}{1+\theta\mu}\hat{s}_{t+T}\right], \tag{A38}$$

$$\hat{p}_t^*(i) = \omega\beta E_t\hat{p}_{t+1}^*(i) + (1-\omega\beta)\left(\hat{p}_t + \frac{1}{1+\theta\mu}\hat{s}_t\right). \tag{A39}$$

All the variables of the right hand side of equation (A38) are aggregate variables, therefore all firms re-optimizing their price choose the exact same price. Let $\hat{p}_t^*$ denote this price. Since the firms who re-optimize their price is randomly determined, the average price of the firms who could not re-optimize their price is equal to the average price level in the previous period plus the inflation rate of the previous period scaled by the degree of indexation. The log-linear consumer price index then simplifies to,

$$\hat{p}_t = (1-\omega)\hat{p}_t^* + \omega\left(\hat{p}_{t-1} + \gamma\pi_{t-1}\right).$$

Solving this for $\hat{p}_t^*$ yields,

$$\hat{p}_t^* = \frac{1}{1-\omega}\left(\hat{p}_t - \omega\hat{p}_{t-1} - \omega\gamma\pi_{t-1}\right) \tag{A40}$$

Substituting (A40) into (A39) to eliminate $\hat{p}_t^*(i)$ and $E_t\hat{p}_{t+1}^*(i)$ leads to the Phillips curve,

$$\pi_t = \left(\frac{1}{1-\beta\gamma}\right)\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega(1+\theta\mu)}\hat{s}_t\right], \tag{A41}$$

where $\pi_t \equiv \hat{p}_t - \hat{p}_{t-1}$.

When there is no capital in the model it is possible to rewrite the Phillips curve in terms of the output gap instead of the marginal cost. When prices are fully flexible, $\omega = 0$, and the objective function for the intermediate firm, given in equation (A25) can be rewritten as,

$$\left(\frac{p_t^f(i)}{p_t^f}\right)^{1-\theta} y_t^f - S\left[\left(\frac{p_t^f(i)}{p_t^f}\right)^{-\theta} y_t^f\right]. \tag{A42}$$

The first order condition imlies,

$$s_t^f(i) = \frac{1-\theta}{\theta}\left(\frac{p_t^f(i)}{p_t^f}\right). \tag{A43}$$

Since prices are fully flexible, $p_t^f(i) = p_t^f$ for all t and the marginal cost remains constant. The expression for the marginal cost given in equation (A18) implies under flexible prices that,

$$\mu\hat{y}_t^f - \hat{\lambda}_t^f - (\mu+1)\hat{z}_t = 0. \tag{A44}$$

Solving this for z_t and substituting this into the Phillips curve, equation (A41), leads to following expression of the Phillips curve in terms of the output gap,

$$\pi_t = \left(\frac{1}{1-\beta\gamma}\right)\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega(1+\theta\mu)}\left(\mu\tilde{y}_t - \tilde{\lambda}_t\right)\right]. \tag{A45}$$

A.2.6 Phillips Curve with Capital

Deriving the Phillips curve when there is firm-specific capital is substantially more complicated. Equation (A36) shows that each firm's optimal price will depend on its capital stock relative to the aggregate capital stock. Since a firm's capital stock is dependent on its entire investment history, the optimal price will depend on the firm's entire history of being able to re-optimize its price. The convenient result from the previous section that each firm will choose the same price does not hold when there is firm-specific capital and Calvo pricing.

Equation (A34) implicitly defines the optimal choice for the price of intermediate good i in terms of expectations of aggregate variables and the following expectation of the firm's future relative capital stocks:

$$E_t\sum_{T=0}^{\infty}(\omega\beta)^T\,\tilde{k}_{t+T}(i) \tag{A46}$$

To derive the Phillips curve, we must rewrite the above expression in terms of the firm's current capital stock, the current optimal price, and expectations of aggregate variables. The optimal choice for $\tilde{k}_{t+1}(i)$ in terms of expected future output is given in equation (A24). Substituting the log-linear demand for $y_{t+1}(i)$ into equation (A24) leads to,

$$\begin{aligned}\phi\left(\tilde{k}_{t+1}(i)-\tilde{k}_t(i)\right) &= \beta\phi\left(E_t\tilde{k}_{t+2}(i)-\tilde{k}_{t+1}(i)\right)\\ &-\left[\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right]\left[\theta(\mu+1)E_t\tilde{p}_{t+1}(i)-(1+\alpha\mu)\tilde{k}_{t+1}(i)\right].\end{aligned} \tag{A47}$$

where $\tilde{p}_{t+1}(i)=\hat{p}_{t+1}(i)-\hat{p}_{t+1}$ is the relative price of intermediate good i in period $t+1$. The rational expectations solution for (A47) must have the form,

$$\tilde{k}_{t+1}(i)=m\tilde{k}_t(i)+n\tilde{p}_t(i), \tag{A48}$$

where m and n are determined by the method of undetermined coefficients in the next subsection. For a firm re-optimizing their price, this equation can be rewritten as

$$\tilde{k}_{t+1}(i)=m\tilde{k}_t(i)+n\hat{p}^*_t(i)-n\hat{p}_t(i). \tag{A49}$$

Substituting this into equation (A46) shows that,

$$E_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\tilde{k}_{t+T+1}(i)=mE_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\tilde{k}_{t+T}(i)+\frac{n}{1-\omega\beta}\hat{p}^*_t(i)-nE_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\hat{p}_{t+T}.$$

Multiply both sides of this equation by $(\omega\beta)$ then add $\tilde{k}_t(i)$ to both sides in order to make the summation on the left hand side identical to the summation on the right hand side. Doing this yields,

$$E_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\tilde{k}_{t+T}(i)=\omega\beta mE_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\tilde{k}_{t+T}(i)+\frac{\omega\beta n}{1-\omega\beta}\hat{p}^*_t(i)-\omega\beta nE_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\hat{p}_{t+T}+\tilde{k}_t(i).$$

Solving this equation yields,

$$E_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\tilde{k}_{t+T}(i)=\frac{1}{1-\omega\beta m}\left[\frac{\omega\beta n}{1-\omega\beta}\hat{p}^*_t(i)+\tilde{k}_t(i)-\omega\beta nE_t\sum_{T=0}^{\infty}\left(\omega\beta\right)^T\hat{p}_{t+T}\right]. \tag{A50}$$

Substituting this into equation (A34) and solving for $\hat{p}^*_t(i)$ leads to the following solution,

$$\hat{p}^*_t(i)=(1-\omega\beta)E_t\sum_{T=0}^{\infty}(\omega\beta)^T\left(\hat{p}_{t+T}-\gamma\pi^*_{t+T}+\nu\hat{s}_{t+T}\right)-\frac{\alpha\nu(\mu+1)(1-\omega\beta)}{(1-\alpha)(1-\omega\beta m)}\tilde{k}_t(i), \tag{A51}$$

where,

$$\nu=\left[1+\frac{\theta(\mu+\alpha)}{1-\alpha}+\frac{\alpha\omega\beta n(\mu+1)}{(1-\alpha)(1-\omega\beta m)}\right].$$

Equation (A51) expresses the optimal price of intermediate good i solely in terms of aggregate variables and the firm's current relative capital stock. Since the capital stock was chosen in the previous period, it is independent of whether or not a firm is currently able to re-optimize its price. Therefore the average capital stock among firms re-optimizing their price is equal to the average capital stock in the economy. This implies that average value for $\tilde{k}_t(i)$ over firms re-optimizing their price is equal to zero. Let $\hat{p}_t^*$ denote the average price among these firms. Equation (A51) implies,

$$\hat{p}_t^* = (1-\omega\beta)E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left(\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \nu\hat{s}_{t+T}\right).$$

This can be rewritten as the first order difference equation,

$$\hat{p}_t^* = \omega\beta E_t \hat{p}_{t+1}^* + (1-\omega\beta)\left(\hat{p}_t + \nu\hat{s}_t\right). \tag{A52}$$

Substituting equation (A40) into (A52) to eliminate $\hat{p}_t^*$ and $E_t\hat{p}_{t+1}^*$ leads to the Phillips curve,

$$\pi_t = \left(\frac{1}{1-\beta\gamma}\right)\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \kappa\hat{s}_t\right], \tag{A53}$$

where,

$$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\nu\omega}.$$

A.2.7 Method of Undetermined Coefficients

This subsection uses the method of undetermined coefficients to compute the values of m and n in equation (A49) which must satisfy the optimality condition for capital given in equation (A47). Equation (A47) can be rearranged as,

$$\tilde{k}_{t+1}(i) = \tilde{k}_t(i) + \beta E_t\tilde{k}_{t+2}(i) - \zeta_0 E_t\tilde{p}_{t+1}(i) - \zeta_1\tilde{k}_{t+1}(i), \tag{A54}$$

where ζ_0 and ζ_1 are given by,

$$\zeta_0 = \frac{\theta\left(\mu+1\right)\left[1-\beta\left(1-\delta\right)\right]}{\phi\left(1-\alpha\right)}$$

$$\zeta_1 = \beta + \frac{\left(1+\alpha\mu\right)\left[1-\beta\left(1-\delta\right)\right]}{\phi\left(1-\alpha\right)}$$

I begin by finding an expression for $E_t\tilde{p}_{t+1}(i)$ in terms of $\tilde{k}_t(i)$ and $\tilde{p}_t(i)$. Using equation (A37), the expected relative price can be rewritten as,

$$E_t\tilde{p}_{t+1}(i) = E_t\hat{p}_{t+1}(i) - E_t\hat{p}_{t+1} = \omega\hat{p}_t(i) + (1-\omega)E_t\hat{p}_{t+1}^*(i) - E_t\hat{p}_{t+1} \tag{A55}$$

In order to express $\tilde{p}_{t+1}(i)$ only in terms of $\tilde{p}_t(i)$ and $\tilde{k}_t(i)$, we must next find a solution for $\hat{p}^*_t(i)$. According to equation (A39), the rational expectation solution for $\hat{p}^*_t(i)$ must take the form,

$$\hat{p}^*_t(i) = f(\hat{p}_t, \hat{s}_t) + a\tilde{k}_t(i), \tag{A56}$$

where $f(\cdot)$ is a linear function of aggregate variables and a needs to be determined by the method of undetermined coefficients. Let $\mathcal{L}_t$ denote the set of firms re-optimizing their price in period t. The average price of the firms who are able to re-optimize their price is given by,

$$\hat{p}^*_t = \frac{1}{1-\omega}\int_{i\in\mathcal{L}_t} \hat{p}^*_t(i)di = f(\hat{p}_t, \hat{s}_t) + \frac{a}{1-\omega}\int_{i\in\mathcal{L}_t} \tilde{k}_t(i)di$$

Since $\tilde{k}_t(i)$ was chosen in period $t-1$, it is independent of whether a firm is re-optimizing its price. Therefore the average difference between a firm's capital stock and the aggregate capital stock among firms re-optimizing their price is equal to zero. Therefore,

$$\hat{p}^*_t = f(\hat{p}_t, \hat{s}_t),$$

and equation (A56) can be rewritten as,

$$\hat{p}^*_t(i) = \hat{p}^*_t + a\tilde{k}_t(i). \tag{A57}$$

Advancing equation (A57) one period and taking expectations yields,

$$E_t\hat{p}^*_{t+1}(i) = E_t\hat{p}^*_{t+1} + a\tilde{k}_{t+1}(i), \tag{A58}$$

where $E_t\hat{p}^*_{t+1}$ is the expected average price over firms that can re-optimize their price next period. This can be rewritten in terms of the expected aggregate price level. Since a fraction ω firms will not be able to change their price next period and the remaining $1-\omega$ firms will have an average price $\hat{p}^*_{t+1}$, the expected price level next period is given by,

$$E_t\hat{p}_{t+1} = \omega\hat{p}_t + (1-\omega)E_t\hat{p}^*_{t+1}. \tag{A59}$$

Solving (A59) for $E_t\hat{p}^*_{t+1}$ and substituting this expression into (A58) leads to,

$$E_t\hat{p}^*_{t+1}(i) = \frac{1}{1-\omega}\left(E_t\hat{p}_{t+1} - \omega\hat{p}_t\right) + a\tilde{k}_{t+1}(i). \tag{A60}$$

Substituting equation (A49) for $\tilde{k}_{t+1}(i)$ yields,

$$E_t\hat{p}^*_{t+1}(i) = \frac{1}{1-\omega}\left(E_t\hat{p}_{t+1} - \omega\hat{p}_t\right) + am\tilde{k}_t(i) - an\tilde{p}_t(i). \tag{A61}$$

Plugging this into equation (A55) leads to an expression for $E_t\tilde{p}_{t+1}(i)$ in terms of $\tilde{p}_t(i)$ and

$\tilde{k}_t(i)$,

$$E_t\tilde{p}_{t+1}(i) = [\omega + (1-\omega)\,an]\,\tilde{p}_t(i) + (1-\omega)\,am\tilde{k}_t(i) \tag{A62}$$

Next, using equation (A49), the expected future capital stock is given by,

$$E_t\tilde{k}_{t+2}(i) = m^2\tilde{k}_t(i) + mn\tilde{p}_t(i) + nE_t\tilde{p}_{t+1}(i) \tag{A63}$$

Substituting equation (A62) into equation (A63) leads to an expression for $E_t\tilde{k}_{t+2}(i)$ in terms of $\tilde{p}_t(i)$ and $\tilde{k}_t(i)$,

$$E_t\tilde{k}_{t+2}(i) = \Big[m^2 + amn\,(1-\omega)\Big]\,\tilde{k}_t(i) + \Big[mn + n\omega + an^2(1-\omega)\Big]\,\tilde{p}_t(i) \tag{A64}$$

Plugging in equations (A62), (A63), and (A49) into (A54) leads to an expression for capital of the form given in equation (A49) where m and n must satisfy, respectively,

$$\beta m^2 + [\beta an(1-\omega) - \zeta_1 - \zeta_0 a(1-\omega) - 1]\,m + 1 = 0, \tag{A65}$$

$$\beta a(1-\omega)n^2 + [\beta m + \beta\omega - \zeta_1 - \zeta_0 a(1-\omega) - 1]\,n - \zeta_0\omega = 0. \tag{A66}$$

All that remains is to find an expression for a, also using the method of undetermined coefficients. Substituting the expression for $E_t\hat{p}^*_{t+1}(i)$ given in equation (A61) into equation (A36) and solving for $\hat{p}^*_t(i)$ yields,

$$\begin{aligned}\hat{p}^*_t(i) = \quad & \frac{1}{1-\omega\beta an}\left(\omega\beta am - \frac{\psi\alpha\,(\mu+1)}{1-\alpha}\right)\tilde{k}_t(i) \\ & + \frac{\omega\beta}{(1-\omega)\,(1-\omega\beta an)}\,(E_t\hat{p}_{t+1} - \omega\hat{p}_t) + \hat{p}_t + \frac{\psi}{1-\omega\beta an}\hat{s}_t,\end{aligned} \tag{A67}$$

which implies a must satisfy the quadratic equation,

$$\omega\beta na^2 + (\omega\beta m - 1)\,a - \frac{\alpha\psi(\mu+1)}{1-\alpha} = 0. \tag{A68}$$

Equations (A65), (A66), and (A68) make up a system of quadratic equations that jointly determine the values for m, n, and a in terms of the parameters of the model. Since this is a system of three quadratic equations, there are potentially eight solutions, but these equations alone do not rule out economically infeasible outcomes. Equations (A48) and (A62) can be rewritten as the following dynamic system:

$$\begin{bmatrix} \tilde{k}_{t+1}(i) \\ E_t\tilde{p}_{t+1}(i) \end{bmatrix} = \begin{bmatrix} m & n \\ \omega + (1-\omega)an & 1-\omega \end{bmatrix} \begin{bmatrix} \tilde{k}_t(i) \\ \tilde{p}_t(i) \end{bmatrix}. \tag{A69}$$

The economically feasible solution for m, n, and a must be consistent with stable means and variances of each firm's relative capital stock and relative price. The system is stable if and only if the eigenvalues of the matrix in equation (A69) are inside the unit circle. The eigenvalues are given by,

$$e_1 = \frac{1}{2}\left(m + \omega + (1-\omega)an + \sqrt{[m + \omega + (1-\omega)an]^2 - 4m\omega}\right)$$

$$e_2 = \frac{1}{2}\left(m + \omega + (1-\omega)an - \sqrt{[m + \omega + (1-\omega)an]^2 - 4m\omega}\right)$$

It is evident from these equations that $e_1 > e_2$. Therefore both eigenvalues will be less than 1 in absolute value if and only if $e_1 < 1$ and $e_2 > -1$. The condition on the first eigenvalue implies,

$$\sqrt{[m + \omega + (1-\omega)an]^2 - 4m\omega} < 2 - m - \omega - (1-\omega)an. \tag{A70}$$

Since the left hand side of the inequality is always positive, the left hand side must also be positive. Therefore, squaring both sides preserves the direction of the inequality. Doing this yields,

$$[m + \omega + (1-\omega)an]^2 - 4m\omega < 4 - 4\,[m + \omega + (1-\omega)an] + 4\,[m + \omega + (1-\omega)an]^2 \tag{A71}$$

This inequality does not preserve the restriction implied in (A70) that the right hand side be positive. Therefore (A70) also implies

$$2 - m - \omega - (1-\omega)an > 0. \tag{A72}$$

The inequalities (A71) and (A72) simplify to, respectively,

$$m < 1 - an \tag{A73}$$

$$m < 1 + (1-\omega)(1-an) \tag{A74}$$

The stability condition for the second eigenvalue is,

$$\sqrt{[m + \omega + (1-\omega)an]^2 - 4m\omega} < 2 + m + \omega + (1-\omega)an,$$

which simplifies to,

$$m > -1 - \frac{1-\omega}{1+\omega}an \tag{A75}$$

Finally, the coefficients m, n, and a can be found by the solving the system of quadratic equations (A65), (A66), and (A68), subject to the inequalities (A73), (A74), and (A75).

A.3 Market clearing

Goods market clearing implies total output of the final good is equal to aggregate consumption plus aggregate investment,

$$y_t = c_t + I_t.$$

Log-linearizing this yields,

$$\hat{y}_t = c_y \hat{c}_t + \delta k_y \hat{I}_t, \tag{A76}$$

where c_y is the steady state consumption to output ratio and k_y is the steady state capital to output ratio. The steady state capital to output ratio is found by combining the steady state first order condition for capital rental, given in equation (A10), and the steady state first order condition for investment, given in equation (A20). Evaluating equation (A10) at the steady state and using the steady state marginal cost, given in equation (A30), yields,

$$\rho = \alpha \frac{\theta - 1}{\theta} \left(\frac{y}{k} \right).$$

Evaluating equation (A20) at the steady state yields,

$$1 = \beta \left(\rho + 1 - \delta \right).$$

Combining these equations to eliminate ρ leads to the following capital to output ratio,

$$k_y = \frac{\beta\alpha(\theta - 1)}{\theta \left(1 - \beta + \beta\delta\right)} \tag{A77}$$

Evaluating the goods market clearing condition, (A76), at the steady state yields the following consumption to output ratio,

$$c_y = 1 - \delta k_y. \tag{A78}$$

References

Bullard, J., and J. Duffy (2004): “Learning and structural change in macroeconomic data,” Federal Reserve Bank of St. Louis Working Paper 2004-016A.

Bullard, J., and S. Eusepi (2005): “Did the great inflation occur despite poicymaker commitment to a Taylor rule?,” Review of Economic Dynamics, 8, 324–359.

Bullard, J., G. W. Evans, and S. Honkapohja (2005): “Near Rational Exuberance,” Federal Reserve Bank of St. Louis Working Paper 2004-025B.

Calvo, G. A. (1983): “Staggered prices in a utility maximizing framework,” Journal of Monetary Economics, 12, 383–398.

Carceles-Poveda, A., and C. Giannitsarou (2005): “Adaptive Learning in Practice,” Working Paper.

Cogley, T., and A. Sbordone (2005): “A search for a structural Phillips curve,” Federal Reserve Bank of New Yord Staff Report no. 203.

de Jong, P. (1989): “Smoothing and Interpolation with the State-Space Model,” Journal of the American Statistical Association, 84, 1085–1088.

Evans, G. W., and S. Honkapohja (2001): Learning and Expectations in Macroeconomics. Princeton University Press.

Fuhrer, J. C. (2000): “Habit formation in consumption and its implications for monetary-policy models,” American Economic Review, 90, 367–390.

Hamilton, J. (1994): Time Series Analysis. Princeton University Press.

Ireland, P. (2004): “A method for taking models to the data,” Journal of Economic Dynamics and Control, 28, 1205–1226.

Ireland, P. N. (2005): “Irrational expectations and econometric practice,” Federal Reserve Bank of Atlanta Working Paper 2003-22.

Johnston, L. D., and S. H. Williamson (2007): “The annual real and nominal GDP for the United States, 1790 - present,” URL: http://eh.net/hmit/gdp/.

Lubik, T., and F. Schorfheide (2004): “Testing for indeterminacy: An application to U.S. monetary policy,” American Economic Review, 94, 190–217.

Mark, N. C. (2005): “Changing monetary policy rules, learning, and real exchange rate dyanamics,” NBER Working Paper 11061.

Milani, F. (2005): “Expectations, learning and macroeconomic persistence,” Working Paper.

Orphanides, A., and J. C. Williams (2005a): “Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations,” Journal of Economic Dynamics and Control, 29, 1927–1950.

——— (2005b): “Inflation scares and forecast-based monetary policy,” Review of Economic Dynamics, 8, 498–527.

Primiceri, G. E. (2006): “Why inflation rose and fell: policymakers’ beliefs and US postwar stabilization policy,” Quarterly Journal of Economics, 121, 867–901.

Roberts, J. M. (1995): “New Keynesian Economics and the Phillips Curve,” Journal of Money, Credit and Banking, 27, 975–984.

Rotemberg, J. (1982): “Sticky Prices in the United States,” Journal of Political Economy, 90, 1187–1211.

Rotemberg, J., and M. Woodford (1997): “An optimization based econometric framework for the evaluation of monetary policy," In: Bernanke, B.S. and J. Rotemberg (Eds.), NBER Macroeconomics Annual. MIT Press.

SIMS, C. (2000): "Solving linear rational expectations models," Unpublished manuscript.

SMETS, F., AND R. WOUTERS (2003): "An estimated stochastic dynamic general equilibrium model of the Euro area," Journal of the European Economic Assocation, 1, 1123–1175.

——— (2005): "Comparing shocks and friction in U.S. and Euro area business cycles: A Bayesian DSGE approach," Journal of Applied Econometrics, 20, 161–183.

TAYLOR, J. (1993): "Discretionary versus policy rules in practice," Carnegie-Rochester Conference Series on Public Policy, 39, 195–214.

WILLIAMS, N. (2005): "Adaptive Learning and Business Cycles," Working paper.

WOODFORD, M. (2003): Interest and prices. Princeton University Press.

——— (2005): "Firm specific capital and the New Keynesian Phillips curve," International Journal of Central Banking, 1, 1–46.

Table 1: Model Without Capital: Learning with RE Initial Conditions

		Case 1		Case 2	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.4221	0.1062	0.4241	0.1216
Inverse IES	σ	0.5152	0.4401	0.5236	0.4865
Phillips Slope	κ	0.0001	0.0002	0.0001	0.0002
Price Indexation	γ	0.9900	0.0634	0.9901	0.0907
MP Persistence	ρ_r	0.9207	0.0207	0.9207	0.0214
MP Output	ψ_y	0.4946	0.1901	0.4949	0.1967
MP Inflation	ψ_π	1.9994	0.0000	1.9995	0.0000
Nat. Rate Pers.	ρ_n	0.8488	0.0684	0.8489	0.0645
Cost Push Pers.	ρ_u	0.0000	0.0692	0.0000	0.0608
Nat. Rate Std. Dev.	σ_n	0.0751	0.0706	0.0736	0.0741
Cost Push Std. Dev.	σ_u	0.0029	0.0002	0.0029	0.0003
MP Std. Dev.	σ_r	0.0030	0.0001	0.0030	0.0001
SS Inflation	π^*	5.9904	1.2374	5.9905	1.2739
Learning Gain	g	–	–	0.0067	0.0070
Log-likelihood			-459.9390		-459.5154
MSE Output Gap			0.6087		0.6061
MSE Inflation			1.3313		1.3269
MSE Fed. Funds Rate			1.6480		1.6519

Table 2: Model Without Capital: Learning Without Using Shocks

		Case 1		Case 3	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.4221	0.1062	0.3027	0.1216
Inverse IES	σ	0.5152	0.4401	0.2251	0.4865
Phillips Slope	κ	0.0001	0.0002	0.0004	0.0002
Price Indexation	γ	0.9900	0.0634	0.9999	0.0907
MP Persistence	ρ_r	0.9207	0.0207	0.9131	0.0214
MP Output	ψ_y	0.4946	0.1901	0.4762	0.1967
MP Inflation	ψ_π	1.9994	0.0000	1.9865	0.0000
Nat. Rate Pers.	ρ_n	0.8488	0.0684	0.8413	0.0645
Cost Push Pers.	ρ_u	0.0000	0.0692	0.0002	0.0608
Nat. Rate Std. Dev.	σ_n	0.0751	0.0706	0.2736	0.0741
Cost Push Std. Dev.	σ_u	0.0029	0.0002	0.0054	0.0003
MP Std. Dev.	σ_r	0.0030	0.0001	0.0030	0.0001
SS Inflation	π^*	5.9904	1.2374	5.9539	1.2739
Learning Gain	g	–	–	0.0042	0.0070
Log-likelihood			-459.9390		-458.8326
MSE Output Gap			0.6087		0.6032
MSE Inflation			1.3313		1.3371
MSE Fed. Funds Rate			1.6480		1.6378

Table 3: Model Without Capital: Learning with Pre-sample Initial Conditions

		Case 1		Case 4	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.4221	0.1062	0.5293	0.1216
Inverse IES	σ	0.5152	0.4401	0.2502	0.4865
Phillips Slope	κ	0.0001	0.0002	0.0064	0.0002
Price Indexation	γ	0.9900	0.0634	0.9989	0.0907
MP Persistence	ρ_r	0.9207	0.0207	0.8454	0.0214
MP Output	ψ_y	0.4946	0.1901	0.3200	0.1967
MP Inflation	ψ_π	1.9994	0.0000	1.5109	0.0000
Nat. Rate Pers.	ρ_n	0.8488	0.0684	0.6810	0.0645
Cost Push Pers.	ρ_u	0.0000	0.0692	0.4419	0.0608
Nat. Rate Std. Dev.	σ_n	0.0751	0.0706	0.5835	0.0741
Cost Push Std. Dev.	σ_u	0.0029	0.0002	0.0086	0.0003
MP Std. Dev.	σ_r	0.0030	0.0001	0.0030	0.0001
SS Inflation	π^*	5.9904	1.2374	5.8862	1.2739
Learning Gain	g	–	–	0.0828	0.0070
Log-likelihood			-459.9390		-573.3274
MSE Output Gap			0.6087		0.7989
MSE Inflation			1.3313		2.7104
MSE Fed. Funds Rate			1.6480		1.7396

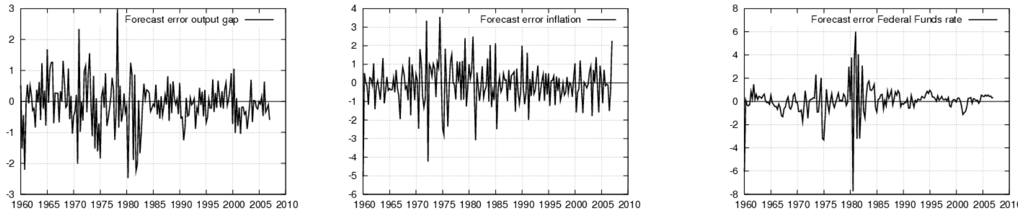
Table 4: Model Without Capital: Learning with Estimated Initial Conditions

		Case 1		Case 5	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.4221	0.1062	0.3052	0.1216
Inverse IES	σ	0.5152	0.4401	0.1960	0.4865
Phillips Slope	κ	0.0001	0.0002	0.0001	0.0002
Price Indexation	γ	0.9900	0.0634	0.9893	0.0907
MP Persistence	ρ_r	0.9207	0.0207	0.9193	0.0214
MP Output	ψ_y	0.4946	0.1901	0.4944	0.1967
MP Inflation	ψ_π	1.9994	0.0000	1.9992	0.0000
Nat. Rate Pers.	ρ_n	0.8488	0.0684	0.8488	0.0645
Cost Push Pers.	ρ_u	0.0000	0.0692	0.0000	0.0608
Nat. Rate Std. Dev.	σ_n	0.0751	0.0706	0.2310	0.0741
Cost Push Std. Dev.	σ_u	0.0029	0.0002	0.0054	0.0003
MP Std. Dev.	σ_r	0.0030	0.0001	0.0030	0.0001
SS Inflation	π^*	5.9904	1.2374	5.9894	1.2739
Learning Gain	g	–	–	0.0000	0.0070
Log-likelihood			-459.9390		-449.3276
MSE Output Gap			0.6087		0.5679
MSE Inflation			1.3313		1.2922
MSE Fed. Funds Rate			1.6480		1.6486

Figure 1: Forecast Errors

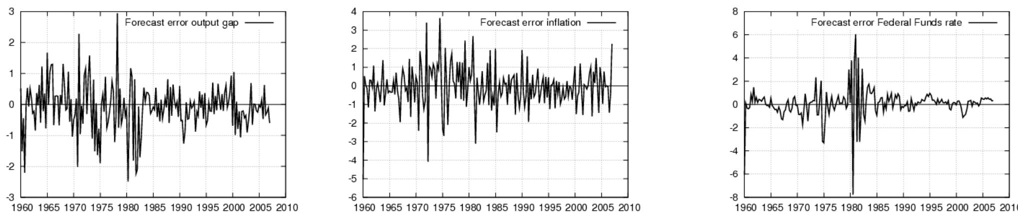
Rational Expectations

Output gap | Inflation | Federal Funds Rate



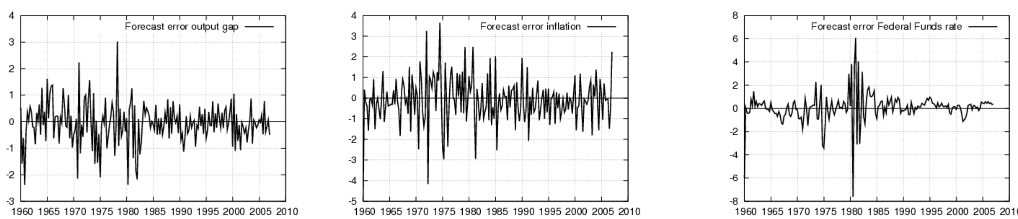
Learning with RE Initial Conditions

Output gap (0.9993) | Inflation (0.9973) | Federal Funds Rate (0.9999)



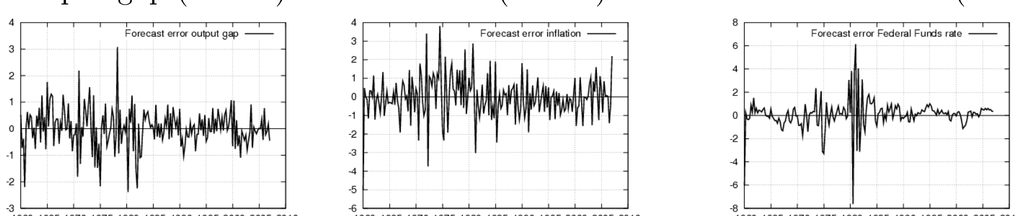
Learning Without Observable Shocks

Output gap (0.9846) | Inflation (0.9960) | Federal Funds Rate (0.9996)



Learning with Pre-sample Initial Conditions

Output gap (0.7344) | Inflation (0.7617) | Federal Funds Rate (0.9481)



Learning with Estimated Initial Conditions

Output gap (0.9682) | Inflation (0.9896) | Federal Funds Rate (0.9997)

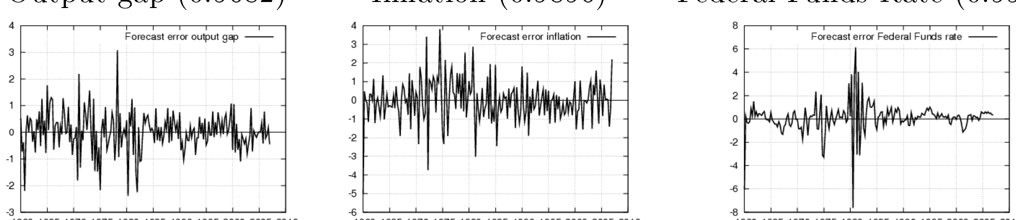
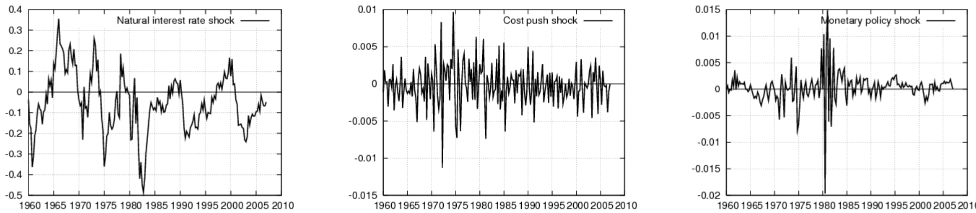


Figure 2: Estimated Shocks

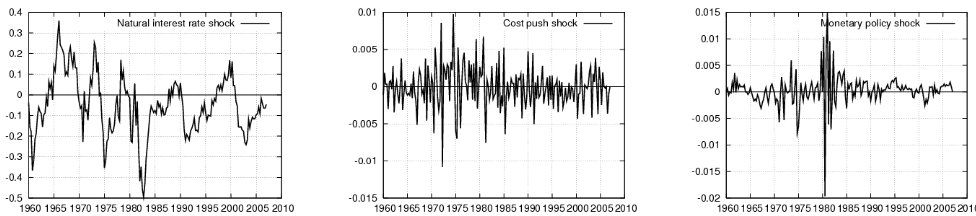
Rational Expectations

Nat. Rate | Cost Push | Policy Shock



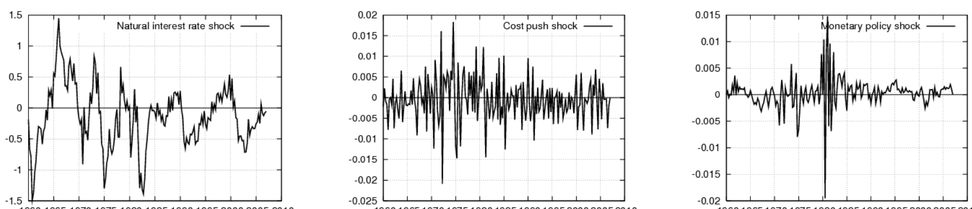
Learning with RE Initial Conditions

Nat. Rate (0.9996) | Cost Push (0.9965) | Policy Shock (1.0000)



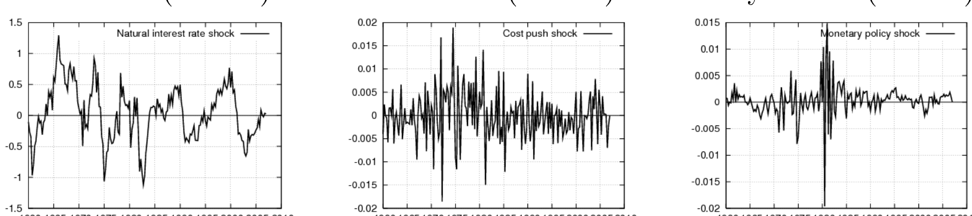
Learning Without Observable Shocks

Nat. Rate (0.9731) | Cost Push (0.9932) | Policy Shock (0.9995)



Learning with Pre-sample Initial Conditions

Nat. Rate (0.3652) | Cost Push (0.6110) | Policy Shock (0.9619)



Learning with Estimated Initial Conditions

Nat. Rate (0.9769) | Cost Push (0.9832) | Policy Shock (0.9998)

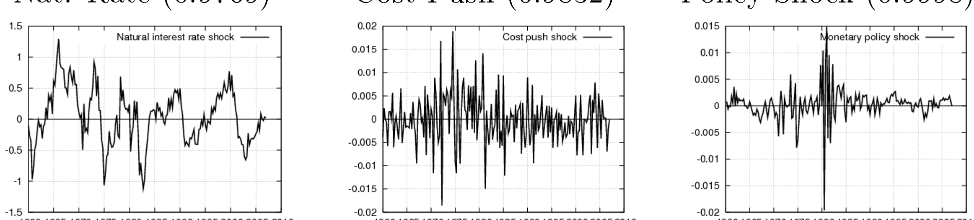
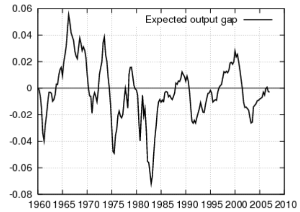
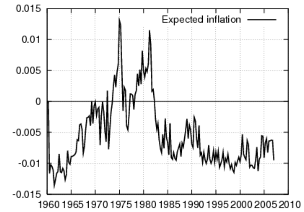


Figure 3: Expectations
Rational Expectations

Output Gap

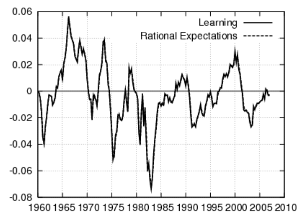


Inflation

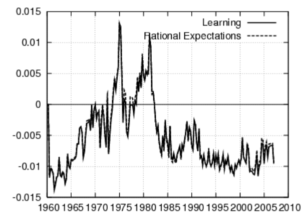


Learning with RE Initial Conditions

Output Gap

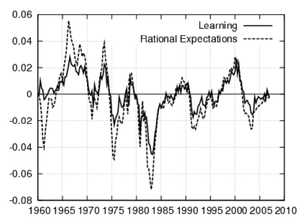


Inflation

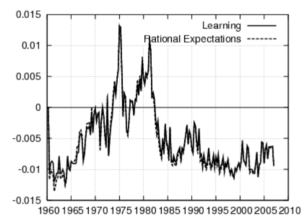


Learning Without Using Shocks

Output Gap

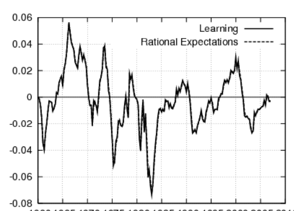


Inflation

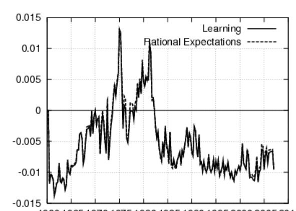


Learning with Pre-sample Initial Conditions

Output Gap

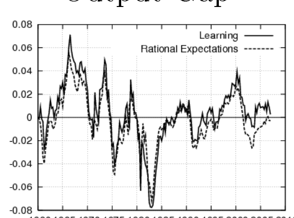


Inflation



Learning with Estimated Initial Conditions

Output Gap



Inflation

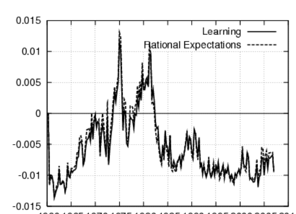


Table 5: Model With Capital: Learning with RE Initial Conditions

		Case 1		Case 2	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.9181	0.1007	0.9181	0.1017
Inverse IES	σ	0.3432	0.7774	0.3432	0.7967
Capital Share	α	0.3584	0.1189	0.3584	0.1219
Cons / Output	c_y	0.8753	0.0044	0.8753	0.0044
Cost Capital Adj.	ϕ	6.9883	1.4836	6.9883	2.8850
Phillips Slope	κ	0.0090	0.0036	0.0090	0.0039
Price Indexation	γ	0.0001	0.0768	0.0001	0.0769
MP Persistence	ρ_r	0.7481	0.0472	0.7481	0.0570
MP Output	ψ_y	0.1003	0.0379	0.1003	0.0379
MP Inflation	ψ_π	1.0014	0.1195	1.0014	0.1219
Tech. Shock Pers.	ρ_z	0.9716	0.0133	0.9716	0.0134
Pref. Shock Pers.	ρ_ξ	0.5647	0.1159	0.5647	0.1160
Inv. Shock Pers.	ρ_μ	0.9050	0.0426	0.9050	0.0435
Tech. Shock Std. Dev.	σ_z	0.0094	0.0041	0.0094	0.0044
Inv. Shock Std. Dev.	σ_μ	0.0306	0.0070	0.0306	0.0099
Pref. Shock Std. Dev.	σ_ξ	0.3587	0.2699	0.3587	0.2741
MP Shock Std. Dev.	σ_r	0.0033	0.0002	0.0033	0.0003
SS Inflation	π^*	0.2209	4.2127	0.2209	4.2253
SS Output (10,000)	y^*	1.4085	0.0212	1.4085	0.0213
Learning gain	g	–	–	0.0000	0.0147
Log-likelihood			-2391.5472		-2391.5472
MSE Consumption			7285.1049		7285.1049
MSE Investment			14454.2922		14454.2922
MSE Inflation			1.2633		1.2633
MSE Fed. Funds Rate			1.7499		1.7499

Table 6: Model With Capital: Learning Without Using Shocks

		Case 1		Case 3	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.9181	0.1007	0.8393	0.1888
Inverse IES	σ	0.3432	0.7774	0.3771	0.8493
Capital Share	α	0.3584	0.1189	0.3870	0.2697
Cons / Output	c_y	0.8753	0.0044	0.8987	0.0000
Cost Capital Adj.	ϕ	6.9883	1.4836	6.9747	2.1739
Phillips Slope	κ	0.0090	0.0036	0.0158	0.0065
Price Indexation	γ	0.0001	0.0768	0.0007	0.0774
MP Persistence	ρ_r	0.7481	0.0472	0.8031	0.0365
MP Output	ψ_y	0.1003	0.0379	0.1005	0.0478
MP Inflation	ψ_π	1.0014	0.1195	1.0285	0.1656
Tech. Shock Pers.	ρ_z	0.9716	0.0133	0.9689	0.0461
Pref. Shock Pers.	ρ_ξ	0.5647	0.1159	0.6644	0.1550
Inv. Shock Pers.	ρ_μ	0.9050	0.0426	0.9182	0.0050
Tech. Shock Std. Dev.	σ_z	0.0094	0.0041	0.1133	0.0708
Inv. Shock Std. Dev.	σ_μ	0.0306	0.0070	0.1026	0.0189
Pref. Shock Std. Dev.	σ_ξ	0.3587	0.2699	0.3532	0.1867
MP Shock Std. Dev.	σ_r	0.0033	0.0002	0.0031	0.0001
SS Inflation	π^*	0.2209	4.2127	3.8982	1.4266
SS Output (10,000)	y^*	1.4085	0.0212	1.4200	0.0181
Learning gain	g	–	–	0.0052	0.0019
Log-likelihood			-2391.5472		-2320.0228
MSE Consumption			7285.1049		9532.5647
MSE Investment			14454.2922		11044.5295
MSE Inflation			1.2633		1.2455
MSE Fed. Funds Rate			1.7499		1.6766

Table 7: Model With Capital: Learning with Pre-sample Initial Conditions

		Case 1		Case 4	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.9181	0.1007	0.7685	0.2555
Inverse IES	σ	0.3432	0.7774	0.6433	1.7938
Capital Share	α	0.3584	0.1189	0.3177	0.3785
Cons / Output	c_y	0.8753	0.0044	0.8944	0.0036
Cost Capital Adj.	ϕ	6.9883	1.4836	6.8234	2.6605
Phillips Slope	κ	0.0090	0.0036	0.0113	0.0035
Price Indexation	γ	0.0001	0.0768	0.0877	0.1071
MP Persistence	ρ_r	0.7481	0.0472	0.8975	0.0382
MP Output	ψ_y	0.1003	0.0379	0.1565	0.1132
MP Inflation	ψ_π	1.0014	0.1195	1.0462	0.1866
Tech. Shock Pers.	ρ_z	0.9716	0.0133	0.8205	0.0357
Pref. Shock Pers.	ρ_ξ	0.5647	0.1159	0.9007	0.0286
Inv. Shock Pers.	ρ_μ	0.9050	0.0426	1.0000	0.0000
Tech. Shock Std. Dev.	σ_z	0.0094	0.0041	0.3079	0.1910
Inv. Shock Std. Dev.	σ_μ	0.0306	0.0070	0.1064	0.0133
Pref. Shock Std. Dev.	σ_ξ	0.3587	0.2699	0.4362	0.4727
MP Shock Std. Dev.	σ_r	0.0033	0.0002	0.0032	0.0001
SS Inflation	π^*	0.2209	4.2127	0.0035	0.9250
SS Output (10,000)	y^*	1.4085	0.0212	1.4074	0.0089
Learning gain	g	–	–	0.0060	0.0012
Log-likelihood			-2391.5472		-2506.8255
MSE Consumption			7285.1049		9584.1202
MSE Investment			14454.2922		44510.7805
MSE Inflation			1.2633		3.5212
MSE Fed. Funds Rate			1.7499		1.5378

Table 8: Model With Capital: Learning with Estimated Initial Conditions

		Case 1		Case 5	
Description	Parameter	Estimate	Std. Dev.	Estimate	Std. Dev.
Habit Formation	η	0.9181	0.1007	0.8564	6.7771
Inverse IES	σ	0.3432	0.7774	0.1667	16.3052
Capital Share	α	0.3584	0.1189	0.3662	0.6600
Cons / Output	c_y	0.8753	0.0044	0.8851	0.0122
Cost Capital Adj.	ϕ	6.9883	1.4836	6.9999	0.0001
Phillips Slope	κ	0.0090	0.0036	0.0287	0.0201
Price Indexation	γ	0.0001	0.0768	0.0002	0.3186
MP Persistence	ρ_r	0.7481	0.0472	0.9136	0.0641
MP Output	ψ_y	0.1003	0.0379	0.1296	0.2032
MP Inflation	ψ_π	1.0014	0.1195	1.0089	0.4514
Tech. Shock Pers.	ρ_z	0.9716	0.0133	0.9582	0.0411
Pref. Shock Pers.	ρ_ξ	0.5647	0.1159	0.1614	0.1110
Inv. Shock Pers.	ρ_μ	0.9050	0.0426	0.9075	0.0902
Tech. Shock Std. Dev.	σ_z	0.0094	0.0041	0.0609	0.0920
Inv. Shock Std. Dev.	σ_μ	0.0306	0.0070	0.0709	0.0220
Pref. Shock Std. Dev.	σ_ξ	0.3587	0.2699	0.0918	0.2118
MP Shock Std. Dev.	σ_r	0.0033	0.0002	0.0030	0.0001
SS Inflation	π^*	0.2209	4.2127	0.2202	4.7265
SS Output (10,000)	y^*	1.4085	0.0212	1.4177	0.0783
Learning gain	g	–	–	0.0005	0.0018
Log-likelihood			-2391.5472		-2237.0404
MSE Consumption			7285.1049		6702.3679
MSE Investment			14454.2922		6304.4836
MSE Inflation			1.2633		1.1815
MSE Fed. Funds Rate			1.7499		1.4896

Figure 4: Forecast Errors

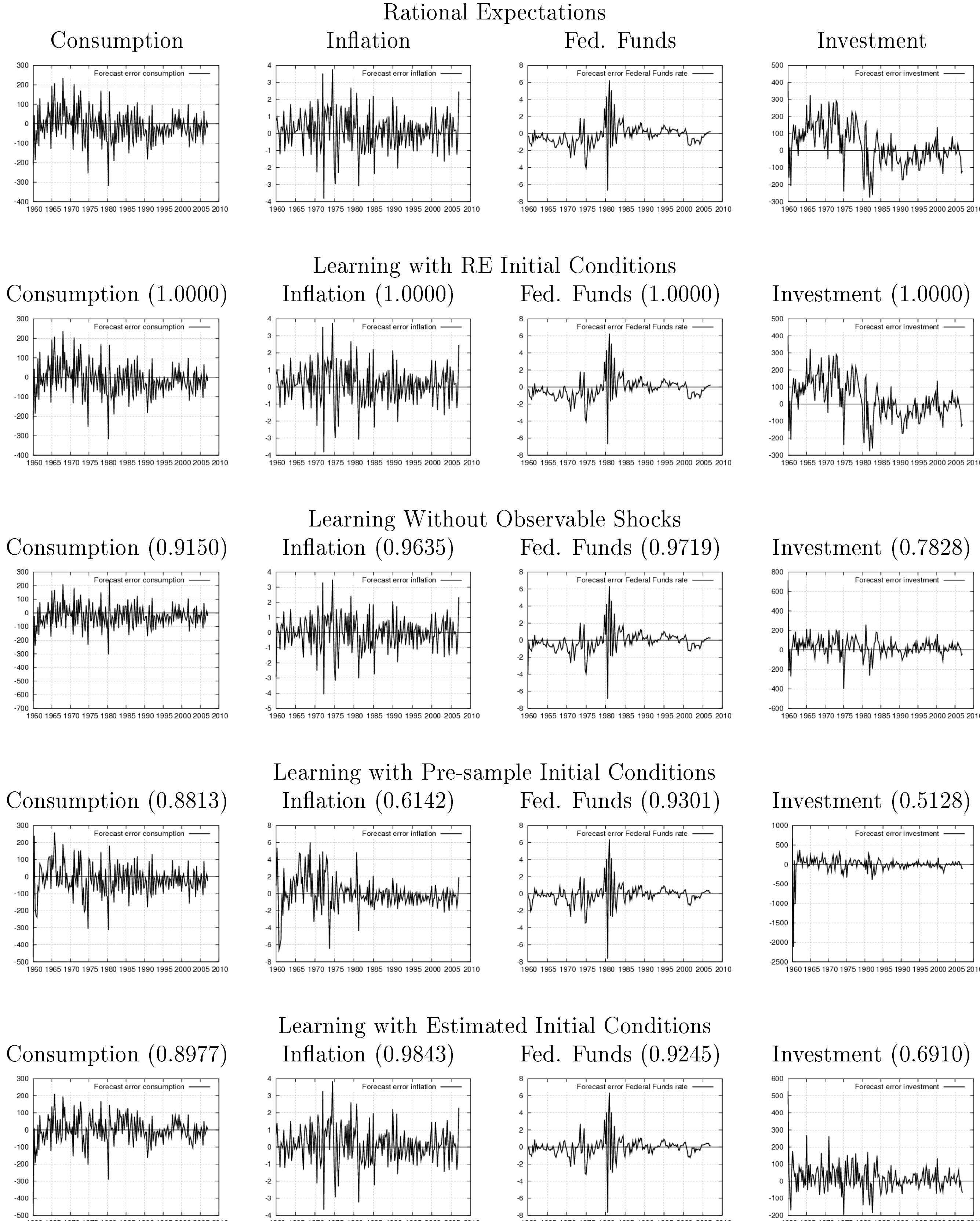


Figure 5: Estimated Shocks

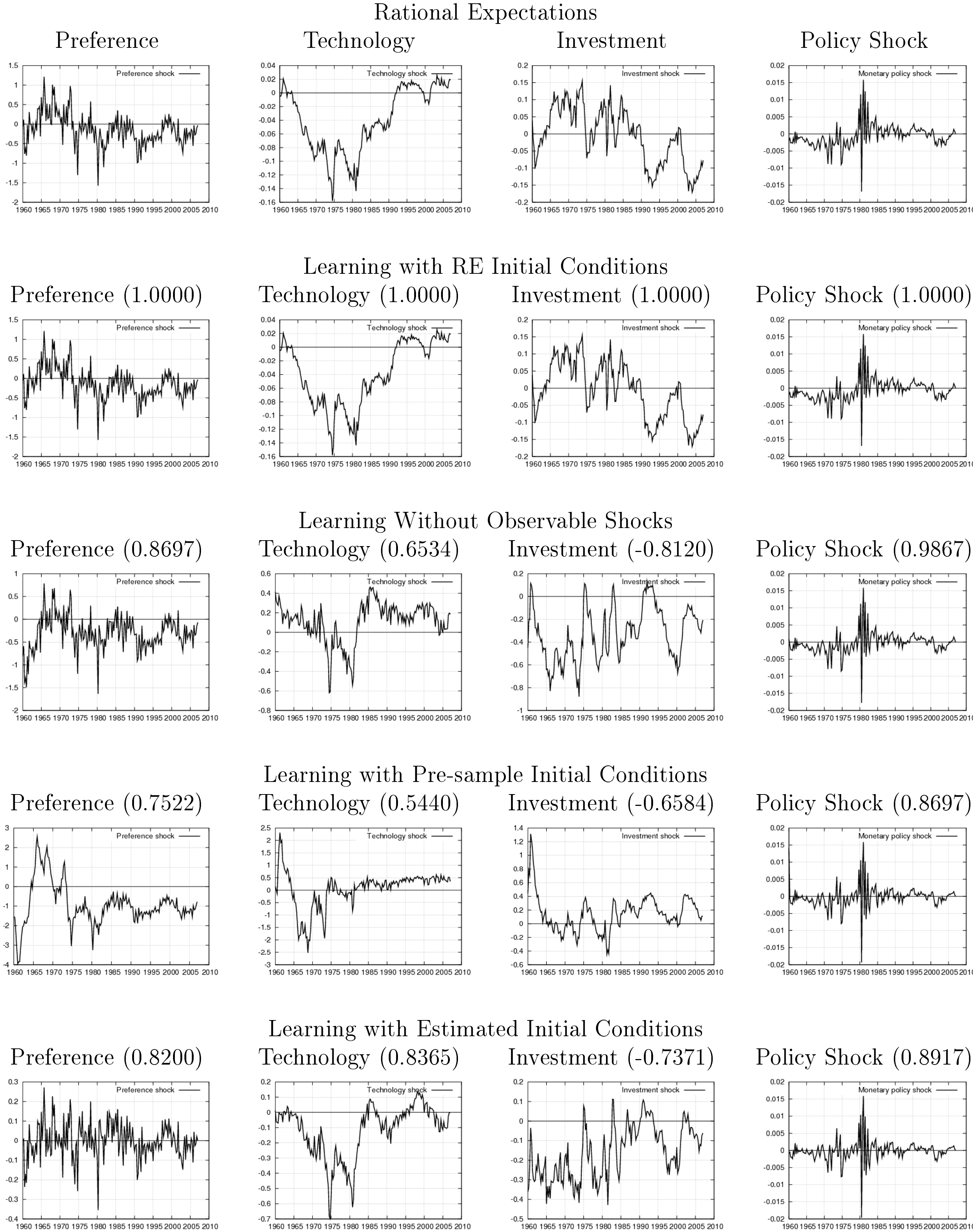
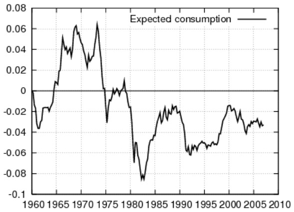


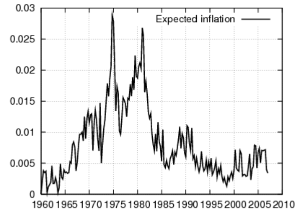
Figure 6: Expectations

Rational Expectations

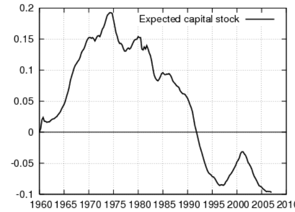
Consumption



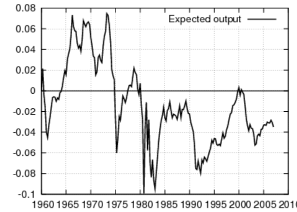
Inflation



Capital Stock

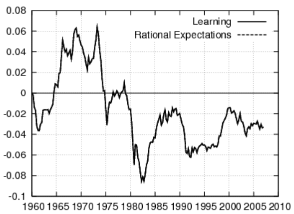


Output

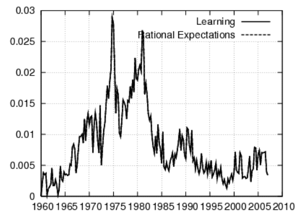


Learning with RE Initial Conditions

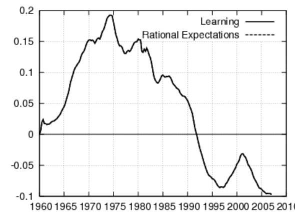
Consumption



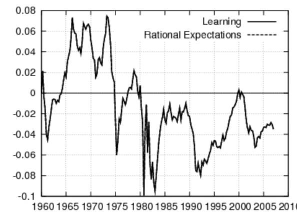
Inflation



Capital Stock

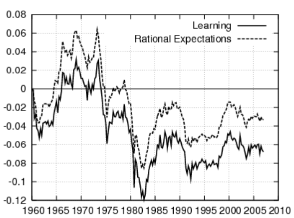


Output

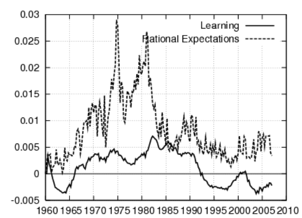


Learning Without Using Shocks

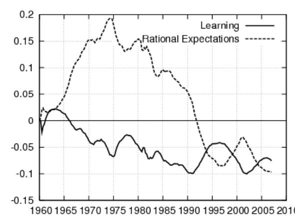
Consumption



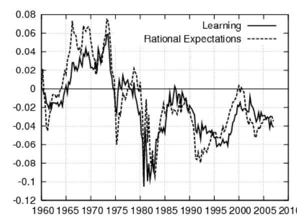
Inflation



Capital Stock

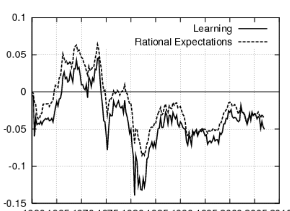


Output

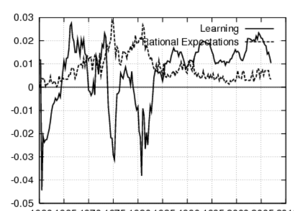


Learning with Pre-sample Initial Conditions

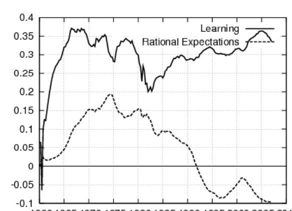
Consumption



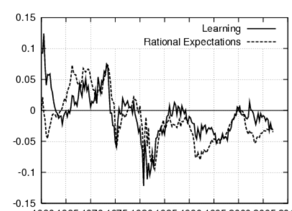
Inflation



Capital Stock

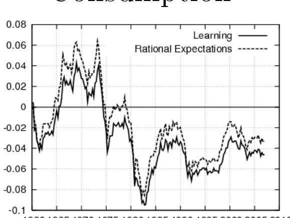


Output

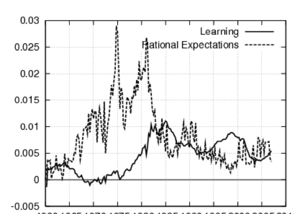


Learning with Estimated Initial Conditions

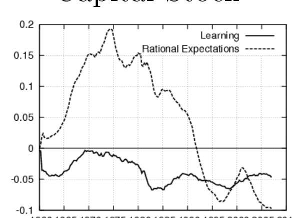
Consumption



Inflation



Capital Stock



Output